Recovering the homology of immersed manifolds

Raphaël Tinarrage

Recovering the homology of immersed manifolds

Raphaël Tinarrage
Datashape, Inria Paris-Saclay – LMO, Université Paris-Saclay

Abstract. Given a sample of an abstract manifold immersed in some Euclidean space, we describe a way to recover the singular homology of the original manifold. It consists in estimating its tangent bundle—seen as subset of another Euclidean space—in a measure theoretic point of view, and in applying measure-based filtrations for persistent homology. The construction we propose is consistent and stable, and does not involve the knowledge of the dimension of the manifold. In order to obtain quantitative results, we introduce the normal reach, which is a notion of reach suitable for an immersed manifold.

Numerical experiments. A Python notebook at `https://github.com/raphaeltinarrage/ImmersedManifolds/blob/master/Demo.ipynb`. Some animations are gathered at `https://youtube.com/playlist?list=PL_FkltNTtklDlIFg1djM5XprlL8YsOhW4`.

Acknowledgements. The author want to thank Frédéric Chazal, Marc Glisse and Théo Lacombe for fruitful discussions and corrections.

Contents

1 Introduction

1.1 Statement of the problem

Let $\mathcal{M}_0$ be a compact $\mathcal{C}^2$-manifold of dimension d, and μ_0 a Radon probability measure on $\mathcal{M}_0$ with support $\mathrm{supp}\,(\mu_0) = \mathcal{M}_0$. Let $E = \mathbb{R}^n$ be the Euclidean space and $u\colon \mathcal{M}_0 \to E$ an immersion. We assume that the immersion is such that self-intersection points correspond to different tangent spaces. In other words, for every $x_0, y_0 \in \mathcal{M}_0$ such that $x_0 \neq y_0$ and $u(x_0) = u(y_0)$, the tangent spaces $d_{x_0}u(T_{x_0}\mathcal{M}_0)$ and $d_{y_0}u(T_{y_0}\mathcal{M}_0)$ are different. Define the image of the immersion $\mathcal{M} = u(\mathcal{M}_0)$ and the pushforward measure $\mu = u_*\mu_0$. We suppose that we are observing the measure μ, or a close measure ν. Our goal is to infer the singular homology of $\mathcal{M}_0$ (with coefficients in $\mathbb{Z}_2$ for instance) from ν.

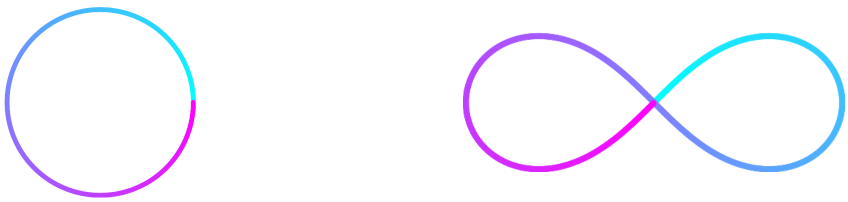

Figure 1: Left: The abstract manifold $\mathcal{M}_0$, diffeomorphic to a circle. Right: The immersion $\mathcal{M} \subset \mathbb{R}^2$, known as the lemniscate of Bernoulli.

As shown in Figure 1, the immersion may self-intersect, hence the singular homology of $\mathcal{M}_0$ and $\mathcal{M}$ may differ. To get back to $\mathcal{M}_0$, we proceed as follows: let $\mathcal{M}(E)$ be the vector space of $n \times n$ matrices, and $\check{u}\colon \mathcal{M}_0 \to E \times \mathcal{M}(E)$ the application

$$\check{u}\colon x_0 \longmapsto \left(u(x_0),\ \frac{1}{d+2} p_{T_x\mathcal{M}} \right),$$

where $p_{T_x\mathcal{M}}$ is the matrix representative of the orthogonal projection on the tangent space $T_x\mathcal{M} \subset E$. Define $\check{\mathcal{M}} = \check{u}(\mathcal{M}_0)$. The set $\check{\mathcal{M}}$ is a submanifold of $E \times \mathcal{M}(E)$, diffeomorphic to $\mathcal{M}_0$. It is called the lift of $\mathcal{M}_0$.

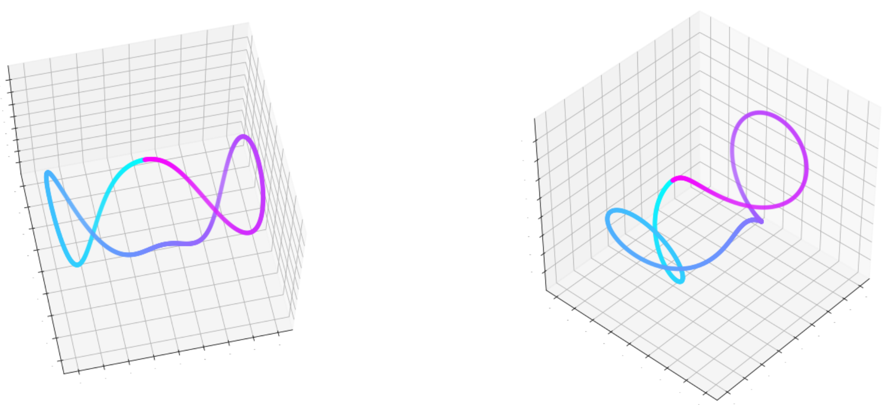

Figure 2: Two views of the submanifold $\check{\mathcal{M}} \subset \mathbb{R}^2 \times \mathcal{M}(\mathbb{R}^2)$, projected in a 3-dimensional subspace via PCA. Observe that it does not self-intersect.

Suppose that one is able to estimate $\check{\mathcal{M}}$ from ν. Then one could consider the persistent homology of a filtration based on $\check{\mathcal{M}}$—say the Čech filtration of $\check{\mathcal{M}}$ in the ambient space $E \times \mathcal{M}(E)$ for instance—and hope to read the singular homology of $\mathcal{M}_0$ in the corresponding persistent barcode.

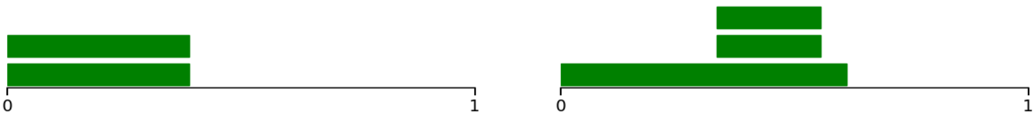

Figure 3: Left: Persistence barcode of the 1-homology of the Čech filtration of $\mathcal{M}$ in the ambient space $\mathbb{R}^2$. One reads the 1-homology of the lemniscate. Right: Persistence barcode of the 1-homology of the Čech filtration of $\check{\mathcal{M}}$ in the ambient space $\mathbb{R}^2 \times \mathcal{M}(\mathbb{R}^2)$. One reads the 1-homology of a circle.

Instead of estimating the lifted submanifold $\check{\mathcal{M}}$, we propose to estimate the *exact lifted measure* $\check{\mu}_0$, defined as $\check{\mu}_0 = \check{u}_* \mu_0$. It is a measure on $E \times \mathcal{M}(E)$, with support $\check{\mathcal{M}}$. Using measure-based filtrations—such as the DTM-filtrations—one can also hope to recover the singular homology of $\mathcal{M}_0$.

It is worth noting that $\check{\mathcal{M}}$ can be naturally seen as a submanifold of $E \times \mathcal{G}_d(E)$, where $\mathcal{G}_d(E)$ denotes the Grassmannian of d-dimensional linear subspaces of E. From this point of view, $\check{\mu}_0$ can be seen as a measure on $E \times \mathcal{G}_d(E)$, i.e., a varifold. However, for computational reasons, we choose to work in $\mathcal{M}(E)$ instead of $\mathcal{G}_d(E)$.

Here is an alternative definition of $\check{\mu}_0$: for any $\phi \colon E \times \mathcal{M}(E) \to \mathbb{R}$ with compact support,

$$\int \phi(x, A) \mathrm{d}\check{\mu}_0(x, A) = \int \phi\left(u(x_0), \frac{1}{d+2} p_{T_x \mathcal{M}}\right) \mathrm{d}\mu_0(x_0).$$

Getting back to the actual observed measure ν, we propose to estimate $\check{\mu}_0$ with the *lifted measure* $\check{\nu}$, defined as follows: for any $\phi \colon E \times \mathcal{M}(E) \to \mathbb{R}$ with compact support,

$$\int \phi(x, A) \mathrm{d}\check{\nu}(x, A) = \int \phi\Big(x, \overline{\Sigma}_\nu(x)\Big) \mathrm{d}\nu(x),$$

where $\overline{\Sigma}_\nu(x)$ is normalized local covariance matrix (defined in Section 3). We prove that $\overline{\Sigma}_\nu(x)$ can be used to estimate the tangent spaces $\frac{1}{d+2} p_{T_x \mathcal{M}}$ of $\mathcal{M}$ (Proposition 3.1), and that it is stable with respect to ν (Equation 26). This estimation may be biased next to multiple points of $\mathcal{M}$, as shown in Figure 4. However, we prove a global estimation result, of the following form: $\check{\mu}_0$ and $\check{\nu}$ are close in the Wasserstein metric, as long as μ and ν are (Theorem 3.10). As a consequence, the persistence diagrams of the DTM-filtrations based on $\check{\mu}_0$ and ν are close in bottleneck distance (Corollary 4.5).

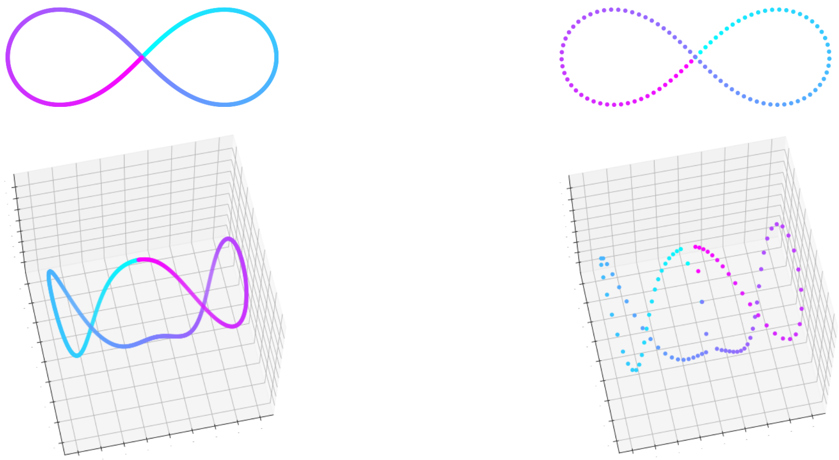

Figure 4: Left: The sets $\operatorname{supp}(\mu) = \mathcal{M}$ and $\operatorname{supp}(\check{\mu}_0) = \check{\mathcal{M}}$, where μ is the uniform measure on the lemniscate. Right: The sets $\operatorname{supp}(\nu)$ and $\operatorname{supp}(\check{\nu})$, where ν is the empirical measure on a 100-sample of the lemniscate. Parameters $\gamma = 2$ and $r = 0{,}1$.

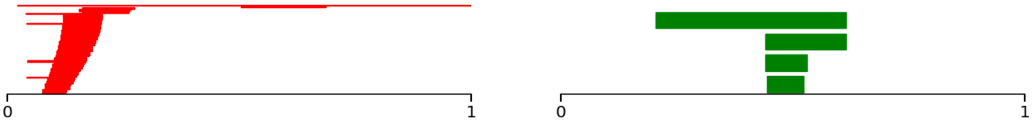

Figure 5: Persistence barcodes of the 0-homology (left) and 1-homology (right) of the DTM-filtration of the lifted measure $\check{\nu}$. Observe that the 1-homology of the circle appears as a large feature of the barcode. Parameters $\gamma = 2$, $r = 0{,}1$ and $m = 0{,}01$.

1.2 Notations and hypotheses

Notations. We adpot the following notations:

- $n, d > 0$ integers.
- If $x, y \in \mathbb{R}$, $x \wedge y$ is the minimum of x and y.
- $E = \mathbb{R}^n$ the Euclidean space, $\mathcal{M}(E)$ the vector space of $n \times n$ matrices, $\mathcal{G}_d(E)$ the Grassmannian.
- A is a subset of E, $\mathrm{med}(A)$ its medial axis, $\mathrm{reach}(A)$ its reach. For every $x \in E$, $\mathrm{dist}(x, A)$ is the distance from x to A.
- For $x, y \in E$, $x \perp y$ denotes the orthogonality of x and y
- If $x, y \in E$, $x \otimes y = x^t y \in \mathcal{M}(E)$ is the outer product, and $x^{\otimes 2} = x \otimes x$.
- $\|\cdot\|$ the Euclidean norm on E and $\langle \cdot, \cdot \rangle$ the corresponding inner product, $\|\cdot\|_{\mathrm{F}}$ the Frobenius norm on $\mathcal{M}(E)$, $\|\cdot\|_\gamma$ the γ-norm on $E \times \mathcal{M}(E)$ (defined in Subsection 3.1).
- $W_p(\cdot, \cdot)$ the p-Wasserstein distance between measures on E, $W_{p,\gamma}(\cdot, \cdot)$ the (p, γ)-Wasserstein distance between measures on $E \times \mathcal{M}(E)$ (defined in Subsection 3.1).
- $\mathcal{H}^d$ the d-dimensional Hausdorff measure on E or on a subspace $T \subset E$.
- If μ is a measure of positive finite mass, $|\mu|$ denotes its mass, $\overline{\mu} = \frac{1}{|\mu|}\mu$ is the associated probability measure, $\check{\mu}$ denotes the associated lifted measure (Subsection 3.1).
- If T is a subspace of E, p_T denotes the orthogonal projection matrix on T.
- $\mathcal{B}(x, r)$ and $\overline{\mathcal{B}}(x, r)$ the open and closed balls of E, $\partial\mathcal{B}(x, r)$ the sphere. V_d and S_{d-1} denote $\mathcal{H}^d(\mathcal{B}(0, 1))$ and $\mathcal{H}^{d-1}(\partial\mathcal{B}(0, 1))$ (note that $S_{d-1} = dV_d$).
- $\mathcal{M}_0$ is a Riemannian manifold, and $\mathcal{B}_{\mathcal{M}_0}(x, r)$ and $\overline{\mathcal{B}}_{\mathcal{M}_0}(x, r)$ denote the open and closed geodesics balls. For $x_0, y_0 \in \mathcal{M}_0$, $\mathrm{d}_{\mathcal{M}_0}(x_0, y_0)$ denotes the geodesic distance.
- If T is a subspace of E, $\mathcal{B}_T(x, r)$ and $\overline{\mathcal{B}}_T(x, r)$ denote the open and closed balls of T for the Euclidean distance .
- if f is a map with values in $\mathbb{R}$ and $t \in \mathbb{R}$, f^t denotes the sublevel set $f^t = f^{-1}((-\infty, t])$.

Model. We consider an abstract $\mathcal{C}^2$-manifold $\mathcal{M}_0$ of dimension d, and an immersion $u\colon \mathcal{M}_0 \to E$. We denote $\mathcal{M} = u(\mathcal{M}_0)$. Moreover, we write $T_{x_0}\mathcal{M}_0$ for the (abstract) tangent space of $\mathcal{M}_0$ at x_0, and $T_x\mathcal{M}$ for $d_{x_0}u(T_{x_0}\mathcal{M}_0)$, which is an affine subspace of E. Let $\check{u}$ be the application

$$\begin{aligned} \check{u}\colon\ \mathcal{M}_0 &\longrightarrow E \times \mathcal{M}(E) \\ x_0 &\longmapsto (x,\ p_{T_x\mathcal{M}})\,, \end{aligned}$$

where $p_{T_x\mathcal{M}}$ is the orthogonal projection matrix on $T_x\mathcal{M}$. We denote $\check{\mathcal{M}} = \check{u}(\mathcal{M}_0)$. We also consider a probability measure μ_0 on $\mathcal{M}_0$, and define $\mu = u_*\mu_0$ and $\check{\mu}_0 = \check{u}_*\mu_0$. These several sets and measures fit in the following commutative diagrams:

$$\begin{tikzcd} \mathcal{M}_0 \arrow[rr, "\check{u}"] \arrow[dr, "u"'] & & \check{\mathcal{M}} \arrow[dl, "\mathrm{proj}"] \\ & \mathcal{M} & \end{tikzcd} \qquad \begin{tikzcd} \mu_0 \arrow[rr, "\check{u}_*"] \arrow[dr, "u_*"'] & & \check{\mu}_0 \arrow[dl, "\mathrm{proj}_*"] \\ & \mu & \end{tikzcd}$$

Moreover, we endow $\mathcal{M}_0$ with the Riemannian structure given by the immersion u. For every $x_0 \in \mathcal{M}_0$, the second fundamental form of $\mathcal{M}_0$ at x_0 is denoted

$$II_{x_0} \colon \ T_{x_0}\mathcal{M}_0 \times T_{x_0}\mathcal{M}_0 \longrightarrow (T_x\mathcal{M})^\perp,$$

and the exponential map is denoted

$$\exp_{x_0}^{\mathcal{M}_0} \colon \ T_{x_0}\mathcal{M}_0 \longrightarrow \mathcal{M}_0.$$

We shall also consider the application $\exp_x^{\mathcal{M}} \colon T_x\mathcal{M} \to \mathcal{M}$, the exponential map seen in $\mathcal{M}$, defined as $u \circ \exp_{x_0}^{\mathcal{M}_0} \circ (d_{x_0}u)^{-1}$.

Notation convention. In the rest of this paper, symbols with 0 as a subscript shall refer to quantities associated to $\mathcal{M}_0$. For instance, a point of $\mathcal{M}_0$ may be denoted x_0, and a curve on $\mathcal{M}_0$ may be denoted γ_0. Symbols with a caron accent shall refer to quantities associated to $\check{\mathcal{M}}$, such as a point $\check{x}$, or a curve $\check{\gamma}$. Symbols with no such subscript or accent shall refer to quantities associated to $\mathcal{M}$, such as x or γ.

In order to simplify the notations, we consider the following convention:

Dropping the 0 subscript to a symbol shall correspond to applying the map u.
Dropping the 0 subscript to a symbol and adding a caron accent shall correspond to applying the map $\check{u}$.

For instance, if x_0 is a point of $\mathcal{M}_0$, then x represents $u(x_0)$, and $\check{x}$ represents $\check{u}(x_0)$. Note that it is possible to have $x = y$ but $T_x\mathcal{M} \neq T_y\mathcal{M}$. Similarly, if $\gamma_0 \colon I \to \mathcal{M}_0$ is a map, then γ represents $u \circ \gamma$, and $\check{\gamma}$ represents $\check{u} \circ \gamma$.

Hypotheses. We shall refer to the following hypotheses:

Hypothesis 1. For every $x_0, y_0 \in \mathcal{M}_0$ such that $x_0 \neq y_0$ and $x = y$, we have $T_x\mathcal{M} \neq T_y\mathcal{M}$.

Hypothesis 2. The operator norm of the second fundamental form of $\mathcal{M}_0$ at each point is bounded by $\rho > 0$.

Hypothesis 3. The measure μ_0 admits a density f_0 on $\mathcal{M}_0$. Moreover, f_0 is L_0-Lipschitz (with respect to the geodesic distance) and bounded by $f_{\min}, f_{\max} > 0$.

Note that Hypothesis 1 ensures that $\check{u}$ is injective, hence that the set $\check{\mathcal{M}}$ is a submanifold of $E \times \mathcal{M}(E)$. The manifolds $\mathcal{M}_0$ and $\check{\mathcal{M}}$ are $\mathcal{C}^1$-diffeomorphic via $\check{u}$. Hypothesis 2 implies the following key property: if $\gamma_0 \colon I \to \mathcal{M}_0$ is an arc-length parametrized geodesic of class $\mathcal{C}^2$, then for all $\forall t \in I$ we have $\|\ddot{\gamma}(t)\| \leq \rho$ (see Equation 1 in Subsection 2.2). Last, in Hypothesis 3, we consider that $\mathcal{M}_0$ is endowed with the natural Hausdorff measure $\mathcal{H}^d_{\mathcal{M}_0}$, obtained by pulling back the d-dimensional Hausdorff measure $\mathcal{H}^d$ on E via the immersion u.

In Subsection 2.3, we define an application $\lambda_0 \colon \mathcal{M}_0 \to \mathbb{R}^+$, called the normal reach. The notation λ_0^r refers to the sublevel set $\lambda_0^{-1}([0, r])$. We consider the following hypothesis:

Hypothesis 4. There exists $c_4 \geq 0$ and $r_4 > 0$ such that, for every $r \in [0, r_4)$, $\mu_0(\lambda_0^r) \leq c_4 r$.

The author thinks that this hypothesis is a consequence of Hypotheses 1, 2 and 3, but has not been able to prove it yet. As a partial result, we prove that it holds when the dimension of $\mathcal{M}_0$ is 1 (Proposition 2.22).

1.3 Background on persistent homology

In the following, we consider interleavings of filtrations, interleavings of persistence modules and their associated pseudo-distances. Their definitions, restricted to the setting of the paper, are briefly recalled in this section. Let $T = \mathbb{R}^+$ and $E = \mathbb{R}^n$ endowed with the standard Euclidean norm.

Filtrations of sets and simplicial complexes. A family of subsets $V = (V^t)_{t\in T}$ of E is a *filtration* if it is non-decreasing for the inclusion, i.e. for any $s, t \in T$, if $s \leq t$ then $V^s \subseteq V^t$. Given $\epsilon \geq 0$, two filtrations $V = (V^t)_{t\in T}$ and $W = (W^t)_{t\in T}$ of E are *ϵ-interleaved* if, for every $t \in T$, $V^t \subseteq W^{t+\epsilon}$ and $W^t \subseteq V^{t+\epsilon}$. The interleaving pseudo-distance between V and W is defined as the infimum of such ϵ:

$$d_i(V, W) = \inf\{\epsilon, V \text{ and } W \text{ are } \epsilon\text{-interleaved}\}.$$

Persistence modules. Let k be a field. A *persistence module* $\mathbb{V}$ over $T = \mathbb{R}^+$ is a pair $\mathbb{V} = ((\mathbb{V}^t)_{t\in T}, (v_s^t)_{s\leq t\in T})$ where $(\mathbb{V}^t)_{t\in T}$ is a family of k-vector spaces, and $(v_s^t \colon \mathbb{V}^s \to \mathbb{V}^t)_{s\leq t\in T}$ a family of linear maps such that:

- for every $t \in T$, $v_t^t \colon V^t \to V^t$ is the identity map,
- for every $r, s, t \in T$ such that $r \leq s \leq t$, we have $v_s^t \circ v_r^s = v_r^t$.

Given $\epsilon \geq 0$, an *ϵ-morphism* between two persistence modules $\mathbb{V}$ and $\mathbb{W}$ is a family of linear maps $(\phi_t \colon \mathbb{V}^t \to \mathbb{W}^{t+\epsilon})_{t\in T}$ such that the following diagram commutes for every $s \leq t \in T$:

$$\begin{tikzcd}
\mathbb{V}^s \arrow[r, "v_s^t"] \arrow[d, "\phi_s"] & \mathbb{V}^t \arrow[d, "\phi_t"] \\
\mathbb{W}^{s+\epsilon} \arrow[r, "w_{s+\epsilon}^{t+\epsilon}"] & \mathbb{W}^{t+\epsilon}
\end{tikzcd}$$

An *ϵ-interleaving* between two persistence modules $\mathbb{V}$ and $\mathbb{W}$ is a pair of ϵ-morphisms $(\phi_t \colon \mathbb{V}^t \to \mathbb{W}^{t+\epsilon})_{t\in T}$ and $(\psi_t \colon \mathbb{W}^t \to \mathbb{V}^{t+\epsilon})_{t\in T}$ such that the following diagrams commute for every $t \in T$:

$$\begin{tikzcd}
\mathbb{V}^t \arrow[rr, "v_t^{t+2\epsilon}"] \arrow[dr, "\phi_t"'] & & \mathbb{V}^{t+2\epsilon} \\
& \mathbb{W}^{t+\epsilon} \arrow[ur, "\psi_{t+\epsilon}"'] &
\end{tikzcd}
\qquad
\begin{tikzcd}
& \mathbb{V}^{t+\epsilon} \arrow[dr, "\phi_{t+\epsilon}"] & \\
\mathbb{W}^t \arrow[ur, "\psi_t"] \arrow[rr, "w_t^{t+2\epsilon}"] & & \mathbb{W}^{t+2\epsilon}
\end{tikzcd}$$

The interleaving pseudo-distance between $\mathbb{V}$ and $\mathbb{W}$ is defined as

$$d_i(\mathbb{V}, \mathbb{W}) = \inf\{\epsilon \geq 0, \mathbb{V} \text{ and } \mathbb{W} \text{ are } \epsilon\text{-interleaved}\}.$$

A persistence module $\mathbb{V}$ is said to be *q-tame* if for every $s, t \in T$ such that $s < t$, the map v_s^t is of finite rank. The q-tameness of a persistence module ensures that we can define a notion of persistence diagram [CdSGO16]. Moreover, given two q-tame persistence modules $\mathbb{V}, \mathbb{W}$ with persistence diagrams $D(\mathbb{V}), D(\mathbb{W})$, the so-called isometry theorem states that $d_b(D(\mathbb{V}), D(\mathbb{W})) = d_i(\mathbb{V}, \mathbb{W})$, where $d_b(\cdot, \cdot)$ denotes the bottleneck distance between diagrams [CdSGO16, Theorem 4.11].

Relation between filtrations and persistence modules. Applying the homology functor to a filtration gives rise to a persistence module where the linear maps between homology groups are induced by the inclusion maps between sets. As a consequence, if two filtrations are ϵ-interleaved then their associated homology persistence modules are also ϵ-interleaved, the interleaving homomorphisms being induced by the interleaving inclusion maps. Moreover, if the modules are q-tame, then the bottleneck distance between their persistence diagrams is upper-bounded by ϵ.

1.4 Background on persistent homology for measures

In this subsection we define the distance to measure (DTM), based on [CCSM11], and the DTM-filtrations, based on [ACG+18]. Let $T = \mathbb{R}^+$ and $E = \mathbb{R}^n$ endowed with the standard Euclidean norm.

Wasserstein distances. Given two probability measures μ and ν over E, a transport plan between μ and ν is a probability measure π over $E \times E$ whose marginals are μ and ν. Let $p \geq 1$. The *p-Wasserstein distance between μ and ν* is defined as

$$W_p(\mu, \nu) = \left(\inf_\pi \int_{E \times E} \|x - y\|^p d\pi(x, y) \right)^{\frac{1}{p}},$$

where the infimum is taken over all the transport plans π. If q is such that $p \leq q$, then an application of Jensen's inequality shows that $W_p(\mu, \nu) \leq W_q(\mu, \nu)$.

DTM. Let μ be a probability measure over E, and $m \in [0, 1)$ a parameter. For every $x \in E$, let $\delta_{\mu,m}$ be the function defined on E by $\delta_{\mu,m}(x) = \inf\left\{r \geq 0, \mu\left(\overline{\mathcal{B}}(x, r)\right) > m\right\}$. The DTM associated to μ with parameter m is the function $d_{\mu,m} \colon E \to \mathbb{R}$ defined as:

$$d^2_{\mu,m}(x) = \frac{1}{m} \int_0^m \delta^2_{\mu,t}(x) dt.$$

When m is fixed and there is no risk of confusion, we may write d_μ instead of $d_{\mu,m}$. We cite two important properties of the DTM:

Proposition 1.1 ([CCSM11, Corollary 3.7])**.** *For every probability measure μ and $m \in [0, 1)$, $d_{\mu,m}$ is 1-Lipschitz.*

Theorem 1.2 ([CCSM11, Theorem 3.5])**.** *Let μ, ν be two probability measures, and $m \in (0, 1)$. Then $\|d_{\mu,m} - d_{\nu,m}\|_\infty \leq m^{-\frac{1}{2}} W_2(\mu, \nu)$.*

The following theorem shows that the sublevel sets $d^t_{\mu,m}$ of $d_{\mu,m}$ can be used to estimate the homotopy type of $\mathrm{supp}(\mu)$.

Theorem 1.3 ([CCSM11, Corollary 4.11])**.** *Let $m \in (0, 1)$, μ any measure on E, and denote $K = \mathrm{supp}(\mu)$. Suppose that $\mathrm{reach}(K) = \tau > 0$, and that μ satisfies the following hypothesis for $r < \left(\frac{m}{a}\right)^{\frac{1}{d}}$: $\forall x \in K, \mu(\mathcal{B}(x, r)) \geq ar^d$. Let ν be another measure, and denote $w = W_2(\mu, \nu)$. Suppose that $w \leq m^{\frac{1}{2}}\left(\frac{\tau}{9} - \left(\frac{m}{a}\right)^{\frac{1}{d}}\right)$. Define $\epsilon = \left(\frac{m}{a}\right)^{\frac{1}{d}} + m^{-\frac{1}{2}} w$ and choose $t \in [4\epsilon, \tau - 3\epsilon]$. Then $d^t_{\mu,m}$ and K are homotopic equivalent.*

DTM-filtrations. We still consider a probability measure μ on E and a parameter $m \in [0, 1)$. For every $t \in T$, consider the set

$$W^t[\mu] = \bigcup_{x \in \mathrm{supp}(\mu)} \overline{\mathcal{B}}\left(x, (t - d_{\mu,m}(x))^+\right),$$

where $\overline{\mathcal{B}}(x, r^+)$ denotes the closed ball of center x and of radius r if $r \geq 0$, or denotes the empty set if $r < 0$. The family $W[\mu] = (W^t[\mu])_{t \geq 0}$ is a filtration of E. It is called the DTM-filtration with parameters $(\mu, m, 1)$. By applying the singular homology functor, we obtain a persistence module, denoted $\mathbb{W}[\mu]$. If $\mathrm{supp}(\mu)$ is bounded, then $\mathbb{W}[\mu]$ is q-tame.

We close this subsection with a stability result for the DTM-filtrations. First, if μ is any measure, define the quantity

$$c(\mu) = \sup_{x \in \mathrm{supp}(\mu)} d_{\mu,m}(x)$$

The term $c(\mu)$ is to be seen as a quantity controling the regularity of μ. In particular, if μ is the uniform measure on a submanifold, it goes to 0 as m does, as shown by the following lemma.

Lemma 1.4. *Suppose that μ satisfies the following for $r < \left(\frac{m}{a}\right)^{\frac{1}{d}}$: $\forall x \in \operatorname{supp}(\mu)$, $\mu(\mathcal{B}(x,r)) \geq ar^d$. Then $c(\mu) \leq c_{1.4} m^{\frac{1}{d}}$ with $c_{1.4} = a^{-\frac{1}{d}}$.*

Theorem 1.5 ([ACG+18, Theorem 4.5])**.** *Consider two probability measures μ, ν on E with supports X and Y. Let μ', ν' be two probability measures with compact supports Γ and Ω such that $\Gamma \subseteq X$ and $\Omega \subseteq Y$. We have*

$$d_i(W[\mu], W[\nu]) \leq m^{-\frac{1}{2}} W_2(\mu, \mu') + m^{-\frac{1}{2}} W_2(\mu', \nu') + m^{-\frac{1}{2}} W_2(\nu', \nu) + c(\mu') + c(\nu').$$

We can restate Theorem 1.5 without mentioning the intermediate measures μ' and ν'. The proof is given in Appendix A.

Corollary 1.6. *Let μ, ν with $W_2(\mu, \nu) = w \leq \frac{1}{4}$. Suppose that μ satisfies the following for $r < \left(\frac{m}{a}\right)^{\frac{1}{d}}$: $\forall x \in \operatorname{supp}(\mu), \mu(\mathcal{B}(x,r)) \geq ar^d$. Then*

$$d_i\left(W[\mu], W[\nu]\right) \leq c_{1.6} \left(\frac{w}{m}\right)^{\frac{1}{2}} + 2c_{1.4} m^{\frac{1}{d}}$$

with $c_{1.6} = 8\operatorname{diam}(\operatorname{supp}(\mu)) + 5$.

2 Reach of an immersed manifold

In this section, we introduce a new notion of reach, adapted to the immersed manifolds. We start by reviewing known facts about the reach.

2.1 Background on reach

Let us recall the definition of the reach of a subset $A \subseteq E$, as done in [Fed59, Definition 4.1]. Let $x \in E \mapsto \operatorname{dist}(x, A) = \inf_{a \in A} \|x - a\|$ be the distance function to A. First, the medial axis of A is defined as

$$\operatorname{med}(A) = \{x \in E,\ \exists a, b \in A \ \text{ s.t. } \ a \neq b \ \text{ and } \ \|x - a\| = \|x - b\| = \operatorname{dist}(x, A)\}.$$

In other words, $\operatorname{med}(A)$ is the set of points $x \in E$ that admit at least two distinct projections on A.

Definition 2.1. Let $a \in A$. The *reach* of A at a (or *local feature size*) is defined as $\operatorname{reach}(A, a) = \operatorname{dist}(a, \operatorname{med}(A))$. The *reach* of A is defined as $\operatorname{reach}(A) = \inf_{a \in A} \operatorname{reach}(A, a)$.

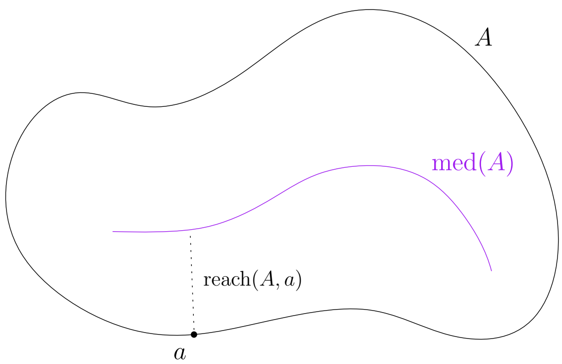

Figure 6: Medial axis and reach of a submanifold of $\mathbb{R}^2$.

In the context of Topological Data Analysis, the reach is a key quantity. For instance, if A is closed subset with positive reach, then for every $t \in [0, \operatorname{reach}(A))$, the t-thickening of A, denoted

A^t, deform retracts on A. Besides, if B is any other subset of E with Hausdorff distance not greater than ϵ from A, then for any $t \in [4\epsilon, \mathrm{reach}(A) - 3\epsilon)$, the thickening B^t deforms retracts on A [CCSL09, Theorem 4.6, case $\mu = 1$]. Consequently, the thickenings of B allow to recover the homology of A.

Among the other properties of a set A with positive reach, a useful one is the approximation by tangent spaces. For a general set A, we define the tangent cone at $x \in A$ as:

$$\mathrm{Tan}(A, x) = \{0\} \cup \left\{ v \in E, \forall \epsilon > 0, \exists y \in A \text{ s.t. } y \neq x, \|y - x\| < \epsilon, \left\| \frac{v}{\|v\|} - \frac{y - x}{\|y - x\|} \right\| < \epsilon \right\}.$$

Note that if A is a submanifold, we recover the usual notion of tangent space.

Theorem 2.1 ([Fed59, Theorem 4.18(2)]). *A closed set $A \subseteq E$ has positive reach τ if and only if for every $x, y \in A$,*

$$\mathrm{dist}\left(y - x, \mathrm{Tan}(A, x)\right) \leq \frac{1}{2\tau} \|y - x\|^2 .$$

Using this property, it is shown in [ACLZ17] that if $A = \mathcal{M}$ is a submanifold with positive reach, one can estimate the tangent spaces of $\mathcal{M}$ via its local covariance matrices. The quality of the estimation depends on $\mathrm{reach}(\mathcal{M})$. However, in our case, the immersion $u\colon \mathcal{M}_0 \to \mathcal{M}$ may be non-injective, and the set $\mathcal{M}$ may be of reach 0. We solve this issue in Subsection 2.3 by introducing the normal reach.

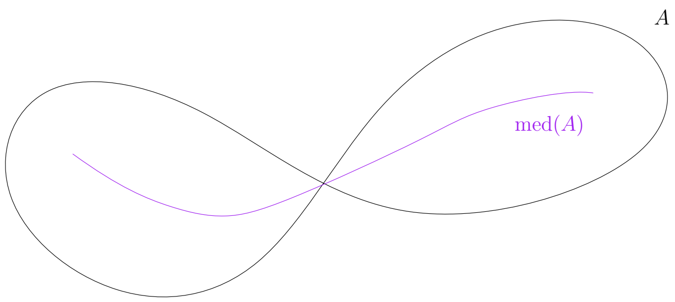

Figure 7: A subset of $\mathbb{R}^2$ with zero reach.

The reach is a quantity that controls both the local and global regularity of the set A. When $A = \mathcal{M}$ is a compact submanifold, it can be shown that $\mathrm{reach}(\mathcal{M})$ is caused either by a bottleneck structure or by high curvature:

Theorem 2.2 ([AKC+19, Theorem 3.4]). *A closed submanifold $\mathcal{M}$ with positive reach must satisfies at least one of the following two properties:*

- Global case: *there exist $x, y \in \mathcal{M}$ with $\|x - y\| = 2\mathrm{reach}(\mathcal{M})$ and $\frac{1}{2}(x + y) \in \mathrm{med}(\mathcal{M})$,*
- Local case: *there exists an arc-length parametrized geodesic $\gamma\colon I \to \mathcal{M}$ with $\|\ddot{\gamma}(0)\| = \mathrm{reach}(\mathcal{M})^{-1}$.*

2.2 Geodesic bounds under curvature conditions

Before introducing the normal reach, we inspect some technical consequences of Hypothesis 2 that shall be used in the rest of the paper.

We consider the immersion $u\colon \mathcal{M}_0 \to \mathcal{M} \subset E$ as in Subsection 1.2. The manifold $\mathcal{M}_0$ is equipped with the Riemannian structure induced by u. For every $x_0 \in \mathcal{M}_0$, the second fundamental form at x_0 is denoted

$$II_{x_0}\colon\ T_{x_0}\mathcal{M}_0 \times T_{x_0}\mathcal{M}_0 \longrightarrow (T_x\mathcal{M})^{\perp}.$$

Let $x_0 \in \mathcal{M}_0$ and consider an arc-length parametrized geodesic $\gamma_0 \colon I \to \mathcal{M}_0$ such that $\gamma_0(0) = x_0$ and $\dot{\gamma}(0) = v_0$. The following relation can be found in [NSW08, Section 6] or [BLW19, Section 3]:

$$II_{x_0}(v_0, v_0) = \ddot{\gamma}_0(0).$$

According to Hypothesis 2, the operator norm of II_{x_0} is bounded by ρ. We deduce that

$$\|\ddot{\gamma}_0(0)\| \leq \rho. \tag{1}$$

Denoting $\gamma = u \circ \gamma_0$, we also have $\|\ddot{\gamma}(0)\| \leq \rho$.

The following lemma is based on this observation. Its second point can be seen as an equivalent of Theorem 2.1, where the Euclidean distance is replaced with the geodesic distance on $\mathcal{M}_0$, and where the quantity $\frac{1}{\rho}$ plays the role of the reach of $\mathcal{M}$.

Lemma 2.3. *Let $x_0 \in \mathcal{M}_0$ and $\gamma_0 \colon I \to \mathcal{M}_0$ an arc-length parametrized geodesic starting from x_0. Let $\gamma = u \circ \gamma_0$ and $v = \dot{\gamma}(0)$. For all $t \in I$, we have*

- $\|\gamma(t) - (x + tv)\| \leq \frac{\rho}{2} t^2$.

As a consequence, for every $y_0 \in \mathcal{M}_0$, denoting $\delta = \mathrm{d}_{\mathcal{M}_0}(x_0, y_0)$, we have

- $\mathrm{dist}\,(y - x, T_x\mathcal{M}) \leq \frac{\rho}{2}\delta^2$,
- $(1 - \frac{\rho}{2}\delta)\delta \leq \|x - y\|$.

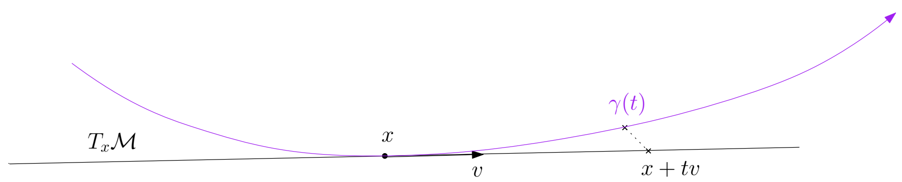

Figure 8: Deviation of a geodesic from its initial direction.

Proof. Consider the application $f \colon t \mapsto \|\gamma(t) - (x + tv)\|$. Since γ is a geodesic, it is of class $\mathcal{C}^2$, and Equation 1 gives $\sup_I \|\ddot{\gamma}\| \leq \rho$. We can apply Taylor-Lagrange formula to get $f(t) \leq \sup_I \|\ddot{\gamma}\| \frac{1}{2} t^2 \leq \frac{\rho}{2} t^2$. Therefore, for all $t \in I$, we have $\|\gamma(t) - (x + tv)\| \leq \frac{\rho}{2} t^2$, and the first claim is proven.

Next, let $\delta = \mathrm{d}_{\mathcal{M}_0}(x_0, y_0)$. By Hopf-Rinow Theorem [dC92, Theorem 2.8 p146], there exists a length-minimizing geodesic γ_0 from x_0 to y_0. Using the last inequality for $t = \delta$ yields

$$\|y - (x + \delta v)\| = \|\gamma(\delta) - (x + \delta v)\| \leq \frac{\rho}{2}\delta^2,$$

and we deduce that $\mathrm{dist}\,(y - x, T_x\mathcal{M}) \leq \|(y - x) - \delta v\| \leq \frac{\rho}{2}\delta^2$.

We prove the last point by applying the triangular inequality:

$$\|x - y\| \geq \|x - (x + \delta v)\| - \|(x + \delta v) - y\| \geq \delta - \frac{\rho}{2}\delta^2. \qquad \square$$

Remark 2.4. The last point of Lemma 2.3 implies the following fact: for all $x_0 \in \mathcal{M}_0$, the map u is injective on the open (geodesic) ball $\mathcal{B}_{\mathcal{M}_0}\left(x_0, \frac{2}{\rho}\right)$. Indeed, if $x_0, y_0 \in \mathcal{M}_0$ are such that $\delta = \mathrm{d}_{\mathcal{M}_0}(x_0, y_0) < \frac{2}{\rho}$, we get $0 < (1 - \frac{\rho}{2}\delta)\delta \leq \|x - y\|$, hence $x \neq y$.

Remark 2.5. We can also deduce the following: for every $y_0 \in \mathcal{B}_{\mathcal{M}_0}\left(x_0, \frac{1}{\rho}\right)$ such that $y_0 \neq x_0$, the vector $y - x$ is not orthogonal to $T_x\mathcal{M}$ nor $T_y\mathcal{M}$. To see this, notice that the inequality $\delta < \frac{1}{\rho}$ and the second point of Lemma 2.3 yields

$$\mathrm{dist}\,(y - x, T_x\mathcal{M}) \leq \frac{\rho}{2}\delta^2 < \frac{1}{2}\delta.$$

Besides, the third point gives $\delta < 2\|y-x\|$, and we deduce that $\operatorname{dist}(y-x, T_x\mathcal{M}) < \|y-x\|$. Equivalently, $y-x$ is not orthogonal to $T_x\mathcal{M}$. Similarly, one proves that $y-x$ is not orthogonal to $T_y\mathcal{M}$.

Consider two points $x_0, y_0 \in \mathcal{M}_0$. We wish to compare their geodesic distance $\mathrm{d}_{\mathcal{M}_0}(x_0, y_0)$ and their Euclidean distance $\|y-x\|$. A first inequality is true in general:

$$\|y-x\| \leq \mathrm{d}_{\mathcal{M}_0}(x_0, y_0).$$

Moreover, if they are close enough in geodesic distance—say $\mathrm{d}_{\mathcal{M}_0}(x_0, y_0) \leq \frac{1}{\rho}$ for instance—then Lemma 2.3 third point yields

$$\mathrm{d}_{\mathcal{M}_0}(x_0, y_0) \leq 2\|x-y\|.$$

However, without any assumption on $\mathrm{d}_{\mathcal{M}_0}(x_0, y_0)$, such an inequality does not hold in general. Figure 9 represents a pair of points which are close in Euclidean distance, but far away with respect to the geodesic distance. In the next subsection, we prove an inequality of the form $\mathrm{d}_{\mathcal{M}_0}(x_0, y_0) \leq c\|x-y\|$, but imposing a constraint on $\|x-y\|$ instead of $\mathrm{d}_{\mathcal{M}_0}(x_0, y_0)$ (see Lemma 2.10).

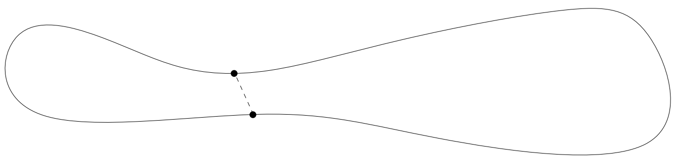

Figure 9: Pair of points for which the geodesic distance is large compared to the Euclidean distance.

We now state a technical lemma. It gives how much time it takes for a geodesic to exit a ball. Its proof is deferred to Appendix B.

Lemma 2.6. *Let $x_0, y_0 \in \mathcal{M}_0$ and $\gamma_0 \colon I \mapsto \mathcal{M}_0$ an arc-length parametrized geodesic with $\gamma_0(0) = y_0$. Define $v = \dot\gamma(0)$. Define $l = \|y-x\|$, and let r be such that $l \leq r < \frac{1}{\rho}$. Consider the application $\phi \colon t \in I \mapsto \|\gamma(t) - x\|^2$.*

- *If $\langle v, y-x\rangle \geq 0$, then $\phi > \phi(0)$ on $(0, T_1)$, where $T_1 = \frac{2}{\rho}\sqrt{1-\rho l}$.*
- *If $\langle v, y-x\rangle = 0$, then ϕ is increasing on $[0, T_2]$ where $T_2 = \frac{\sqrt{2}}{\rho}\sqrt{2-\sqrt{3+\rho^2 l^2}}$.*

Let b be the first value of t such that $\|\gamma(t) - x\| = r$.

- *For all $t \in [0, b]$, we have $\ddot\phi(t) \geq 2(1-\rho r)$.*
- *If $\langle v, y-x\rangle \leq 0$, then $b \geq (1+\rho r)^{-\frac{1}{2}}\sqrt{r^2 - l^2}$.*
- *If $\langle v, y-x\rangle \geq 0$, then $b \leq \left(\frac{1-\rho r}{2}\right)^{-\frac{1}{2}}\sqrt{r^2-l^2}$. Note that if $r < \frac{1}{2\rho}$, then $b < 2r < \frac{1}{\rho}$.*

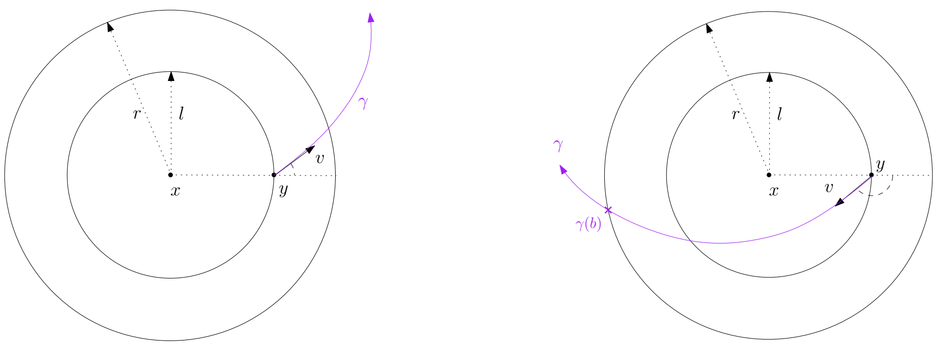

Figure 10: Illustration of Lemma 2.6 first point (left) and fourth point (right).

We close this subsection by studying the exponential map of $\mathcal{M}_0$, denoted

$$\exp_{x_0}^{\mathcal{M}_0}: T_{x_0}\mathcal{M}_0 \to \mathcal{M}_0.$$

According to [AB06, Corollary 4, Point 1], the map $\exp_{x_0}^{\mathcal{M}_0}$ is injective on the open ball $\mathcal{B}_{T_{x_0}\mathcal{M}_0}\left(0, \frac{\pi}{\rho}\right)$ of $T_{x_0}\mathcal{M}_0$, and is a diffeomorphism onto its image $\mathcal{B}_{\mathcal{M}_0}\left(0, \frac{\pi}{\rho}\right)$. We also have a quantitative control of its regularity. Let $x_0 \in \mathcal{M}_0$ and $v_0 \in T_{x_0}\mathcal{M}_0$. The d-dimensional Jacobian of $\exp_{x_0}^{\mathcal{M}_0}$ at v_0 is defined as

$$J_{v_0} = \sqrt{\det\left(A^t A\right)},$$

where $A = d_{v_0}\exp_{x_0}^{\mathcal{M}_0}$ is the differential of the exponential map, seen as a $d \times n$ matrix.

Lemma 2.7. *If $\|v\| = r < \frac{\pi}{2\sqrt{2}\rho}$, the Jacobian J_v of $\exp_{x_0}^{\mathcal{M}_0}$ at v satisfies*

$$\left(1 - \frac{(r\rho)^2}{6}\right)^d \leq J_v \leq \left(1 + (r\rho)^2\right)^d.$$

Proof. The proof is almost identical to [Aam18, Proposition III.22]. From the Gauss equation [dC92, Theorem 2.5 p 130], we get that the sectional curvature $K(v,w)$ of $\mathcal{M}_0$, with v and w orthonormal vectors in $T_{x_0}\mathcal{M}_0$, satisfies

$$K(v,w) = \langle II_{x_0}(v,v), II_{x_0}(w,w)\rangle - \|II_{x_0}(v,w)\|^2.$$

Using Hypothesis 2, we obtain

$$-2\rho^2 \leq K(v,w) \leq \rho^2.$$

Now, let $v \in T_{x_0}\mathcal{M}_0$ and $w \in T_v\left(T_{x_0}\mathcal{M}_0\right) \simeq T_{x_0}\mathcal{M}_0$. As a consequence of the Rauch theorem [DVW15, Lemma 8], the differential of $\exp_{x_0}^{\mathcal{M}_0}$ at v admits the following bound:

$$\left(1 - \frac{(\rho\|v\|)^2}{6}\right)\|w\| \leq \left\|d_v \exp_{x_0}^{\mathcal{M}_0}(w)\right\| \leq \left(1 + (\rho\|v\|)^2\right)\|w\|.$$

Next, denote $A = d_v\exp_{x_0}^{\mathcal{M}_0}$, the differential of the exponential map seen as a $d \times n$ matrix. The last inequality shows that the eigenvalues λ of A^tA are bounded by

$$\left(1 - \frac{(\rho\|v\|)^2}{6}\right)^2 \leq \lambda \leq \left(1 + (\rho\|v\|)^2\right)^2.$$

Since $\det(A^tA)$ is the product of its d eigenvalues, we obtain the result. □

2.3 Normal reach

We still consider an immersion $u: \mathcal{M}_0 \to \mathcal{M} \subset E$ which satisfies Hypothesis 2.

Definition 2.2. For every $x_0 \in \mathcal{M}_0$, let $\Lambda(x_0) = \{y_0 \in \mathcal{M}_0, y_0 \neq x_0, x - y \perp T_y\mathcal{M}\}$. The *normal reach* of $\mathcal{M}_0$ at x_0 is defined as:

$$\lambda_0(x_0) = \inf_{y_0 \in \Lambda(x_0)} \|x - y\|.$$

Observe that if x_0, y_0 are distinct points of $\mathcal{M}_0$ with $x = y$, then $x - y$ is orthogonal to any vector, hence $\lambda_0(x_0) = \|x - y\| = 0$.

Moreover, note that $\Lambda(x_0)$ is closed, hence the infimum of Definition 2.2 is attained. Indeed, we can write $\Lambda(x_0) = L \setminus \{x_0\}$, with $L = \{y_0 \in \mathcal{M}_0, x - y \perp T_y\mathcal{M}\}$. L is a closed set since it is the preimage of $\{0\}$ by the continuous map $y_0 \mapsto \left\|p_{T_y\mathcal{M}}(x-y)\right\|$. Furthermore, $\{x_0\}$ is an isolated point of $\Lambda(x_0)$, since Remark 2.5 says that, for every y_0 in the geodesic ball $\mathcal{B}_{\mathcal{M}_0}\left(x_0, \frac{1}{\rho}\right)$ such that $y_0 \neq x_0$, the vector $x-y$ is not orthogonal to $T_y\mathcal{M}$, hence $y_0 \notin L$.

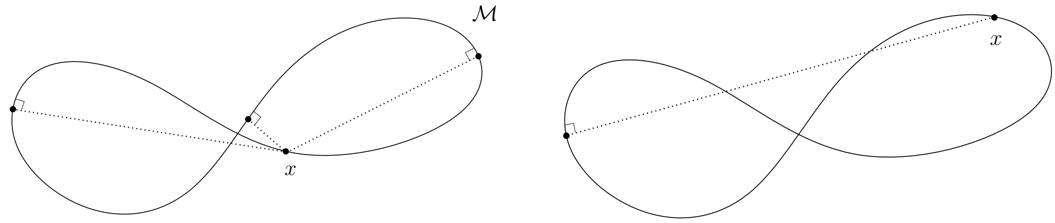

Figure 11: The set $\Lambda(x_0)$ from Definition 2.2, for two different points x_0.

Observe that if a point $x \in \mathcal{M}$ has several preimages by u, then for all $x_0 \in u^{-1}(\{x\})$, we have $\lambda_0(x_0) = 0$. Hence we can define the *normal reach seen in* $\mathcal{M}$, denoted $\lambda \colon \mathcal{M} \to \mathbb{R}$, as

$$\lambda(x) = \begin{cases} \lambda_0\left(u^{-1}(x)\right) & \text{if } x \text{ has only one preimage,} \\ 0 & \text{else.} \end{cases}$$

It satisfies the relation $\lambda_0 = \lambda \circ u$.

Example 2.8. Suppose that $\mathcal{M}$ is the lemniscate of Bernoulli, with diameter 2. Figure 12 represents the values of the normal reach $\lambda \colon \mathcal{M} \to \mathbb{R}$. Note that λ is not continuous.

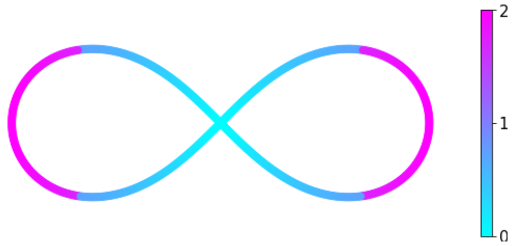

Figure 12: Values of the normal reach on the lemniscate of Bernoulli.

Here is a key property of the normal reach:

Lemma 2.9. *Let $x_0 \in \mathcal{M}_0$. Let $r > 0$ such that $r < \lambda(x)$. Then $u^{-1}\left(\mathcal{M} \cap \overline{\mathcal{B}}(x, r)\right)$ is connected.*

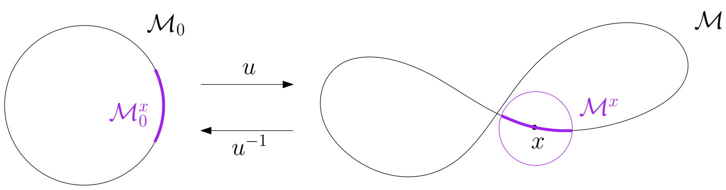

Figure 13: The set $u^{-1}\left(\mathcal{M} \cap \overline{\mathcal{B}}(x, r)\right)$, with $r < \lambda(x)$, is connected.

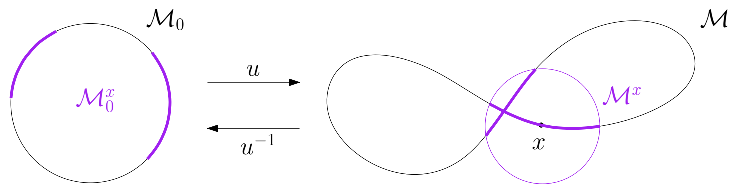

Figure 14: The set $u^{-1}\left(\mathcal{M} \cap \overline{\mathcal{B}}(x,r)\right)$, with $r \geq \lambda(x)$, may not be connected.

Proof. Denote $\mathcal{M}^x = \overline{\mathcal{B}}(x,r) \cap \mathcal{M}$ and $\mathcal{M}_0^x = u^{-1}(\mathcal{M}^x)$. Let us prove that $\mathcal{M}_0^x$ is connected. Suppose that it is not the case. Let $C \subset \mathcal{M}_0^x$ be a connected component which does not contain x_0. Since C is compact, we can consider a minimizer y_0 of $\{\|x-y\|, y_0 \in C\}$. Let us show that $x - y \perp T_y\mathcal{M}$, which will lead to a contradiction.

Two cases may occur: y is in the open ball $\mathcal{B}(x,r)$, or y is on its boundary $\partial\mathcal{B}(x,r)$. If $y \in \mathcal{B}(x,r)$, then there exists a neighborhood $V_0 \subseteq \mathcal{M}_0$ of y such that $V_0 \subseteq \mathcal{M}_0^x$. Hence y satisfies $x - y \perp T_y\mathcal{M}$, otherwise it would not be a local minimizer. Now, suppose that $y \in \partial\mathcal{B}(x,r)$. Since y_0 is a minimizer, there exists a neighborhood $V_0 \subseteq C$ of y_0 such that $V \cap \mathcal{B}(x,r) = \emptyset$. We deduce the existence of a neighborhood $V_0' \subseteq \mathcal{M}_0$ of y_0 such that $V' \cap \mathcal{B}(x,r) = \emptyset$. For instance, take a ball $\mathcal{B} = \mathcal{B}_{\mathcal{M}_0}(y_0,s)$ such that $\mathcal{B} \cap C \subseteq V_0$, and define $V_0' = \mathcal{B}$. We deduce that $y - x \perp T_y\mathcal{M}$.

To conclude, the properties $x - y \perp T_y\mathcal{M}$ and $x_0 \neq y_0$ imply that $\|x-y\| \geq \lambda(x)$, which contradicts $r < \lambda(x)$. □

The following lemma is an equivalent of [NSW08, Proposition 6.3] for the normal reach. It allows to compare the geodesic and Euclidean distance by only imposing a condition on the last one.

Lemma 2.10. *Let $x_0, y_0 \in \mathcal{M}_0$. Denote $r = \|x-y\|$ and $\delta = \mathrm{d}_{\mathcal{M}_0}(x_0,y_0)$. Suppose that $\|x-y\| < \frac{1}{2\rho} \wedge \lambda(x)$. Then*

$$\delta \leq c_{2.10}(\rho r)r \qquad \text{where} \qquad c_{2.10}(t) = \frac{1}{t}\left(1 - \sqrt{1-2t}\right).$$

In other words, the following inclusion holds: $u^{-1}(\overline{\mathcal{B}}(x,r)) \subseteq \overline{\mathcal{B}}_{\mathcal{M}_0}(x_0, c_{2.10}(\rho r)r)$.

Note that, for $t < \frac{1}{2}$, we have the inequalities $1 \leq c_{2.10}(t) \leq 1 + 2t < 2$.

Proof. Denote $\mathcal{M}^x = \overline{\mathcal{B}}(x,r) \cap \mathcal{M}$, $\mathcal{M}_0^x = u^{-1}(\mathcal{M}^x)$ and $\delta = \mathrm{d}_{\mathcal{M}_0}(x,y)$.

Step 1: Let us prove that $\mathcal{M}_0^x \cap \partial\mathcal{B}_{\mathcal{M}_0}(x_0, \delta_{\min} + \epsilon) = \emptyset$, with $\delta_{\min} = c_{2.10}(\rho r)r$, where $c_{2.10}(t) = \frac{1}{t}\left(1 - \sqrt{1-2t}\right)$ and ϵ is small enough. Choose $y_0 \in \partial\mathcal{B}_{\mathcal{M}_0}(x_0, \delta_{\min} + \epsilon)$. According to Lemma 2.3, we have

$$\|x-y\| \geq \left(1 - \frac{\rho}{2}(\delta_{\min} + \epsilon)\right)\left(\delta_{\min} + \epsilon\right). \tag{2}$$

Consider the polynomial $\phi\colon t \mapsto \left(1 - \frac{\rho}{2}t\right)t - r$. Its discriminant is $1 - 2\rho r > 0$, and we deduce that $\phi(t)$ is positive if and only if $t \in \left(\frac{1}{\rho}\left(1 - \sqrt{1-2\rho r}\right), \frac{1}{\rho}\left(1 + \sqrt{1-2\rho r}\right)\right)$. Observe that the first value $\frac{1}{\rho}\left(1 - \sqrt{1-2\rho r}\right)$ is equal to $c_{2.10}(\rho r)r = \delta_{\min}$. Hence $\phi(\delta_{\min} + \epsilon) > 0$ for $0 < \epsilon < \frac{2}{\rho}\sqrt{1-2\rho r}$, and Equation 2 gives $\|x-y\| > r$.

In other words, $y \notin \overline{\mathcal{B}}(x,r)$. This being true for every $y_0 \in \partial\mathcal{B}_{\mathcal{M}_0}(x_0, \delta_{\min} + \epsilon)$, we have $\mathcal{M}_0^x \cap \partial\mathcal{B}_{\mathcal{M}_0}(x_0, \delta_{\min} + \epsilon) = \emptyset$.

Step 2: Let us deduce that $\mathcal{M}_0^x \subseteq \mathcal{B}_{\mathcal{M}_0}(x_0, \delta_{\min})$. By contradiction, if a point $z_0 \in \mathcal{M}_0$ with $\|z - x\| > \delta_{\min}$ were to be in $\mathcal{M}_0^x$, it would be in the connected component of x_0 in $\mathcal{M}_0^x$, since it is connected by Lemma 2.9. But since $\mathcal{M}_0$ is a manifold, this would imply the existence of a continuous path from x_0 to z_0 in $\mathcal{M}_0^x$. But such a path would go through a sphere $\partial\mathcal{B}_{\mathcal{M}_0}(x_0, \delta_{\min} + \epsilon)$, which contradicts Step 1. □

The following proposition connects the normal reach to the usual notion of reach.

Proposition 2.11. *Suppose that $u\colon \mathcal{M}_0 \to \mathcal{M} \subset E$ is an embedding. Let $\tau > 0$ be the reach of $\mathcal{M}$. We have*

$$\tau = \frac{1}{\rho_*} \wedge \frac{1}{2}\lambda_*.$$

where ρ_* is the supremum of the operator norms of the second fundamental forms of $\mathcal{M}_0$, and $\lambda_* = \inf_{x\in\mathcal{M}} \lambda(x)$ is the infimum of the normal reach.

Proof. We first prove that $\tau \geq \frac{1}{\rho_*} \wedge \frac{1}{2}\lambda_*$. According to Theorem 2.2, two cases may occur: the reach is either caused by a bottleneck or by curvature. In the first case, there exists $x, y \in \mathcal{M}$ and $z \in \mathrm{med}(\mathcal{M})$ with $\|x - y\| = 2\tau$ and $\|x - z\| = \|y - z\| = \tau$. We deduce that $x - y \perp T_y\mathcal{M}$. Hence by definition of $\lambda(x)$,

$$\lambda(x) \leq \|x - y\| = 2\|x - z\| \leq 2\tau.$$

In the second case, there exists $x \in \mathcal{M}$ and an arc-length parametrized geodesic $\gamma\colon I \to \mathcal{M}$ such that $\gamma(0) = x$ and $\|\ddot{\gamma}(0)\| = \frac{1}{\tau}$. But $\|\ddot{\gamma}(0)\| \leq \rho_*$, hence $\frac{1}{\tau} \leq \rho_*$.

This disjunction shows that $\tau \geq \frac{1}{\rho} \wedge \frac{1}{2}\lambda_{\min}$.

We now prove that $\tau \leq \frac{1}{\rho_*} \wedge \frac{1}{2}\lambda_*$. The inequality $\tau \leq \frac{1}{\rho_*}$ appears in [NSW08, Proposition 6.1]. To prove $\tau \leq \frac{1}{2}\lambda_*$, consider any $x_0 \in \mathcal{M}_0$. Let $y_0 \in \Lambda(x_0)$ such that $\|x - y\|$ is minimal. Using Theorem 2.1 and the property $x - y \perp T_y\mathcal{M}$, we immediately have

$$\tau \leq \frac{\|x - y\|^2}{2\mathrm{dist}(y - x, T_y\mathcal{M})} = \frac{\|x - y\|}{2} = \frac{\lambda(x)}{2}.$$ □

In the case where u is not an embedding, $\mathcal{M}$ may have zero reach. However, as shown by the following theorem, the normal reach gives a scale at which $\mathcal{M}$ still behaves well. Note that we shall not make use of this result in the rest of the paper.

Theorem 2.12. *Assume that $\mathcal{M}_0$ satisfies Hypothesis 2. Let $x \in \mathcal{M}_0$ and $r < \frac{1}{4\rho} \wedge \lambda(x)$. Then $\overline{\mathcal{B}}(x, r) \cap \mathcal{M}$ is a set of reach at least $\frac{1-2\rho r}{\rho}$. In particular, it is greater than $\frac{1}{2\rho}$.*

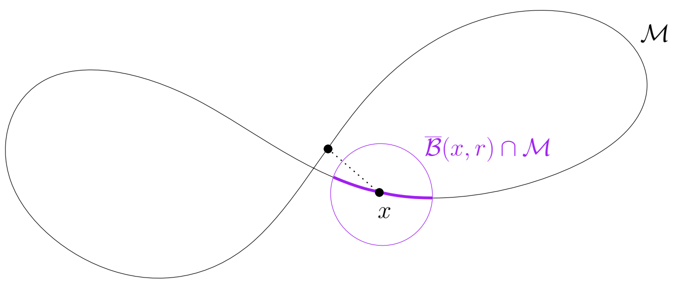

Figure 15: The set $\overline{\mathcal{B}}(x, r) \cap \mathcal{M}$ has positive reach.

Proof. Denote $\mathcal{M}^x = \overline{\mathcal{B}}(x, r) \cap \mathcal{M}$ and $\mathcal{M}_0^x = u^{-1}(\mathcal{M}^x)$.

Step 1: Let us prove that for every $y_0, z_0 \in \mathcal{M}_0^x$,

$$\operatorname{dist}(z-y, T_y\mathcal{M}) \leq \frac{\rho}{2(1-2\rho r)} \|z-y\|^2 .$$

Let $y_0, z_0 \in \mathcal{M}_0^x$, and $\delta = \mathrm{d}_{\mathcal{M}_0}(y_0, z_0)$. Lemma 2.3 Point 3 gives $\delta \leq \frac{1}{1-\frac{\rho}{2}\delta} \|y-z\|$. Moreover, $\delta \leq \mathrm{d}_{\mathcal{M}_0}(y_0, x_0) + \mathrm{d}_{\mathcal{M}_0}(x_0, z_0) \leq 2c_{2.10}(\rho r)r$. Hence,

$$\frac{1}{1-\frac{\rho}{2}\delta} \leq \frac{1}{1-c_{2.10}(\rho r)\rho r} = \frac{1}{\sqrt{1-2\rho r}},$$

and we deduce that

$$\delta \leq \frac{1}{\sqrt{1-2\rho r}} \|y-z\| . \tag{3}$$

Besides, Lemma 2.3 Point 2 gives $\operatorname{dist}(z-y, T_y\mathcal{M}) \leq \frac{\rho}{2}\delta^2$, and combining these two inequalities yields $\operatorname{dist}(z-y, T_y\mathcal{M}) \leq \frac{\rho}{2(1-2\rho r)} \|z-y\|^2$.

Step 2: Let us prove that

$$\operatorname{dist}(z-y, \operatorname{Tan}(\mathcal{M}^x, y)) \leq \frac{\rho}{2(1-2\rho r)} \|z-y\|^2 , \tag{4}$$

where $\operatorname{Tan}(\mathcal{M}^x, y)$ is the tangent cone at y of the closed set $\mathcal{M}^x$.

If $y \in \mathcal{B}(x,r)$, then $\operatorname{Tan}(\mathcal{M}^x, y) = T_y\mathcal{M}$, and the inequality follows from Step 1. Otherwise, suppose that $y \in \partial\mathcal{B}(x,r)$ and that $z \neq y$. Let $\delta = \mathrm{d}_{\mathcal{M}_0}(y_0, z_0)$. According to Equation 3, the inequality $\|y-z\| \leq 2r$ and the assumption $r < \frac{1}{4\rho}$, we have $\delta < \frac{1}{\rho}$. Consider a length-minimizing geodesic $\gamma_0 \colon [0,\delta] \to \mathcal{M}_0$ from y_0 to z_0, and denote $v = \dot\gamma(0)$. Let us show that $v \in \operatorname{Tan}(\mathcal{M}^x, y)$, and we will conclude with Step 1.

Since $\mathcal{M}^x = \overline{\mathcal{B}}(x,r) \cap \mathcal{M}$, $v \in \operatorname{Tan}(\mathcal{M}^x, y)$ is implied by $\langle v, y-x\rangle < 0$. Suppose by contradiction that $\langle v, y-x\rangle \geq 0$. Hence, according to Lemma 2.6 Point 1, with $l = r < \frac{1}{2\rho}$, we have $T_1 = \frac{2}{\rho}\sqrt{1-\rho l} > \frac{\sqrt{2}}{\rho} > \delta$, and

$$\|z-x\| = \|\gamma(\delta)-x\| > \|\gamma(0)-x\| = \|y-x\| = r.$$

We deduce the contradiction $z \notin \overline{\mathcal{B}}(x,r)$.

To conclude the proof, it follows from Theorem 2.1 and Equation 4 that $\mathcal{M}^x$ has reach at least $\frac{1-2\rho r}{\rho}$. $\square$

2.4 Probabilistic bounds under normal reach conditions

We now consider $\mathcal{M}_0$ and μ_0 which satisfy the hypotheses 2 and 3. The aim of this subsection is to provide a quantitative control of the measure $\mu = u_*\mu_0$ (Propositions 2.17 and 2.18). We do so by pulling-back μ on the tangent spaces $T_x\mathcal{M}$, where it is simpler to compute integrals (Lemma 2.15).

Recall that the exponential map of $\mathcal{M}_0$ at a point x_0 is denoted

$$\exp_{x_0}^{\mathcal{M}_0} \colon T_{x_0}\mathcal{M}_0 \to \mathcal{M}_0.$$

To ease the reading of this subsection, we introduce the exponential map seen in $\mathcal{M}$, denoted $\exp_x^{\mathcal{M}} \colon T_x\mathcal{M} \to \mathcal{M}$. It is defined as

$$\exp_x^{\mathcal{M}} = u \circ \exp_{x_0}^{\mathcal{M}_0} \circ (d_{x_0}u)^{-1}.$$

It fits in the following commutative diagram:

$$\begin{tikzcd}
\mathcal{M}_0 \arrow[r, "u"] & \mathcal{M} \\
T_{x_0}\mathcal{M}_0 \arrow[u, "\exp_{x_0}^{\mathcal{M}_0}"] \arrow[r, "d_{x_0}u"] & T_x\mathcal{M} \arrow[u, dashed, "\exp_x^{\mathcal{M}}"']
\end{tikzcd}$$

We also define the map $\overline{\exp}_x^{\mathcal{M}}$ as the restriction of $\exp_x^{\mathcal{M}}$ to the closed ball $\overline{\mathcal{B}}_{T_x\mathcal{M}}\left(0, \frac{\pi}{\rho}\right)$. It is injective by Lemma 2.7. The next lemma gather results of the last subsections. The d-dimensional Jacobian of $\overline{\exp}_x^{\mathcal{M}}$ at v is defined as

$$J_v = \sqrt{\det\left(A^t A\right)},$$

where $A = d_v \overline{\exp}_x^{\mathcal{M}}$ is the differential of the exponential map seen as a $d \times n$ matrix.

Lemma 2.13. *Let $x_0 \in \mathcal{M}_0$ and $r < \frac{1}{2\rho} \wedge \lambda(x)$. Denote $\overline{\mathcal{B}} = \overline{\mathcal{B}}\left(x, r\right)$ and $\overline{\mathcal{B}}^T = \left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}\left(\overline{\mathcal{B}}\right)$. We have the inclusions*

$$\overline{\mathcal{B}}_{T_x\mathcal{M}}\left(0, r\right) \ \subseteq \ \overline{\mathcal{B}}^T \ \subseteq \ \overline{\mathcal{B}}_{T_x\mathcal{M}}\left(0, c_{2.10}(\rho r) r\right).$$

Moreover, for all $v \in \overline{\mathcal{B}}^T$, the d-dimensional Jacobian of $\overline{\exp}_x^{\mathcal{M}}$, denoted J_v, is bounded by

$$\left(1 - \frac{(r\rho)^2}{6}\right)^d \leq J_v \leq \left(1 + (r\rho)^2\right)^d,$$

and these terms are bounded by $J_{\min} = (\frac{23}{24})^d$ and $J_{\max} = (\frac{5}{4})^d$.

Proof. The inclusions come from Lemma 2.10. The bounds on the Jacobian come from Lemma 2.7 and the fact that $c_{2.10}(\rho r) r \leq 2r \leq \frac{1}{\rho} \leq \frac{\pi}{2\sqrt{2}\rho}$ when $r < \frac{1}{2\rho}$. ☐

We now study the measure μ. An application of the coarea formula shows that μ admits the following density against $\mathcal{H}^d_{\mathcal{M}}$, the d-dimensional Hausdorff measure restricted to $\mathcal{M}$:

$$f(x) = \sum_{x_0 \in u^{-1}(\{x\})} f_0(x_0).$$

In particular, if x has only one preimage by u—i.e., if $\lambda\left(x\right) > 0$—then $f(x) = f_0 \circ u^{-1}(x)$. In the rest of the paper, we shall only use f on points x such that $\lambda(x) > 0$.

Remark 2.14. Recall that, by Hypothesis 3, the density f_0 is L_0-Lipschitz with respect to the geodesic distance: for all $x_0, y_0 \in \mathcal{M}_0$,

$$|f_0(x_0) - f_0(y_0)| \leq L_0 \cdot \mathrm{d}_{\mathcal{M}_0}\left(x_0, y_0\right).$$

We can deduce the following: for all $x_0, y_0 \in \mathcal{M}_0$ such that $\|x - y\| < \frac{1}{2\rho} \wedge \lambda(x)$, we have

$$|f(x) - f(y)| \leq L\left\|x - y\right\|$$

with $L = 2L_0$. To prove this, we start with the case where y has only one preimage by u. Since $\|x - y\| < \lambda(x)$ by assumption, we have $0 < \lambda(x)$, hence x also has only one preimage. Now we can write

$$\begin{aligned} |f(x) - f(y)| &= \left|f_0 \circ u^{-1}(x) - f_0 \circ u^{-1}(y)\right| \\ &\leq L_0 \cdot \mathrm{d}_{\mathcal{M}_0}\left(u^{-1}(x), u^{-1}(y)\right) \\ &\leq 2L_0\left\|x - y\right\|, \end{aligned}$$

where we used Lemma 2.10 on the last inequality. Now we prove that $\|x - y\| < \frac{1}{2\rho} \wedge \lambda(x)$ implies that y has only one preimage. Let $r = \|x - y\|$, and suppose by contradiction that y_0, y_1 are two distincts preimages. According to Remark 2.4, $\mathrm{d}_{\mathcal{M}_0}\left(y_0, y_1\right) \geq \frac{2}{\rho}$. But Lemma 2.10 says that $u^{-1}(\mathcal{B}\left(x, r\right)) \subseteq \mathcal{B}_{\mathcal{M}_0}\left(x_0, 2r\right) \subseteq \mathcal{B}_{\mathcal{M}_0}\left(x_0, \frac{1}{\rho}\right)$, which contradicts $\mathrm{d}_{\mathcal{M}_0}\left(y_0, y_1\right) \geq \frac{2}{\rho}$.

Lemma 2.15. *Let $x_0 \in \mathcal{M}_0$ and $r < \frac{1}{2\rho} \wedge \lambda(x)$. Consider μ_x, the measure μ restricted to $\overline{\mathcal{B}}(x, r)$, and define*

$$\nu_x = \left(\overline{\exp}_x^{\mathcal{M}}\right)_*^{-1} \mu_x.$$

The measure ν_x admits the following density against the d-dimensional Hausdorff measure on $T_x\mathcal{M}$:

$$g(v) = f\left(\overline{\exp}_x^{\mathcal{M}}(v)\right) \cdot J_v \cdot 1_{\overline{\mathcal{B}}^T}(v).$$

Moreover, for all $v \in \overline{\mathcal{B}}^T$, the map g satisfies

$$|g(v) - g(0)| \leq c_{2.15} r,$$

where $c_{2.15} = 4L_0 J_{\max} + \frac{d}{2}\rho f_{\max}$.

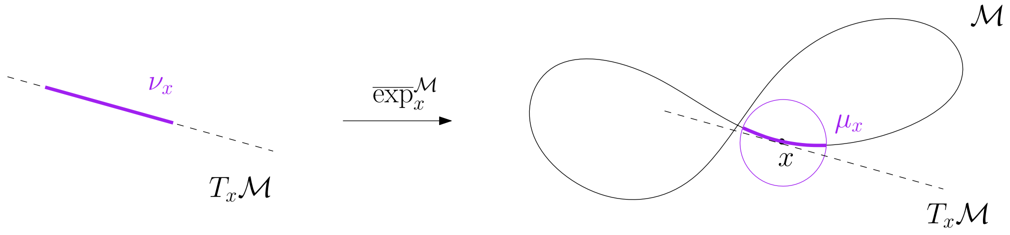

Figure 16: Measures involved in Lemma 2.15.

Proof. The expression of g comes from the area formula [Fed14, Theorem 3.2.5]. To prove the inequality, observe that we can decompose

$$\begin{aligned} g(v) - g(0) &= f\big(\exp_x^{\mathcal{M}}(v)\big) J_v - f\big(\exp_x^{\mathcal{M}}(0)\big) J_0 \\ &= \Big[f\big(\exp_x^{\mathcal{M}}(v)\big) - f\big(\exp_x^{\mathcal{M}}(0)\big) \Big] J_v + (J_v - J_0)\, f\big(\exp_x^{\mathcal{M}}(0)\big) \end{aligned}$$

On the one hand, using Remark 2.14, we get

$$\begin{aligned} \left| f\left(\overline{\exp}_x^{\mathcal{M}}(v)\right) - f\left(\overline{\exp}_x^{\mathcal{M}}(0)\right) \right| &\leq L \left\| \overline{\exp}_x^{\mathcal{M}}(v) - \overline{\exp}_x^{\mathcal{M}}(0) \right\| \\ &= L \left\| u \circ \exp_{x_0}^{\mathcal{M}_0}(v) - u \circ \exp_{x_0}^{\mathcal{M}_0}(0) \right\| \\ &\leq L \cdot \mathrm{d}_{\mathcal{M}_0}\left(\overline{\exp}_{x_0}^{\mathcal{M}_0}(v), x_0 \right) = L \|v\|. \end{aligned}$$

On the other hand, $J_0 = 1$ and $\left(1 - \frac{(r\rho)^2}{6}\right)^d \leq J_v \leq \left(1 + (r\rho)^2\right)^d$ yield $|J_v - J_0| \leq d(\rho r)^2 \leq \frac{d}{2}\rho r$. We eventually obtain

$$g(v) - g(0) \leq L \|v\| J_{\max} + f_{\max} \frac{d}{2} \rho r \leq \left(2L J_{\max} + f_{\max} \frac{d}{2} \rho \right) r. \qquad \square$$

Remark 2.16. In the same vein as Lemma 2.15, define $\overline{\exp}_{x_0}^{\mathcal{M}_0}$ to be the map $\exp_{x_0}^{\mathcal{M}_0}$ restricted to $\overline{\mathcal{B}}_{T_{x_0}\mathcal{M}_0}\left(0, \frac{\pi}{\rho}\right)$. For any $x_0 \in \mathcal{M}_0$, let $\mu_0^{x_0}$ be the measure μ_0 restricted to $\overline{\mathcal{B}}_{\mathcal{M}_0}\left(x_0, \frac{1}{\rho}\right)$, and define the measure

$$\nu_0 = (\overline{\exp}_{x_0}^{\mathcal{M}_0})^{-1} \mu_0^{x_0}.$$

Using the area formula, one shows that ν_0 admits the following density over the d-dimensional Hausdorff measure on $T_{x_0}\mathcal{M}_0$:

$$g_0(v) = f_0\left(\overline{\exp}_{x_0}^{\mathcal{M}_0}(v)\right) \cdot J_v \cdot 1_{\overline{\mathcal{B}}_{T_{x_0}\mathcal{M}_0}\left(0, \frac{1}{\rho}\right)}(v).$$

Now we can use the density g of Lemma 2.15 to derive explicit bounds on μ.

Proposition 2.17. *Let* $x_0 \in \mathcal{M}_0$, $r \leq \frac{1}{2\rho} \wedge \lambda(x)$ *and* $s \in [0, r]$. *We have*

- $\mu\left(\overline{\mathcal{B}}(x, r)\right) \geq c_5 r^d$
- $\left|\frac{\mu(\overline{\mathcal{B}}(x,r))}{V_d r^d} - f(x)\right| \leq c_{2.17} r$
- $\mu\left(\overline{\mathcal{B}}(x, r) \setminus \overline{\mathcal{B}}(x, s)\right) \leq c_6 r^{d-1}(r - s)$

with $c_5 = f_{\min} J_{\min} V_d$, $c_{2.17} = c_{2.15} + f_{\max} J_{\max} d 2^d \rho$ *and* $c_6 = d 2^d f_{\max} J_{\max} V_d$.

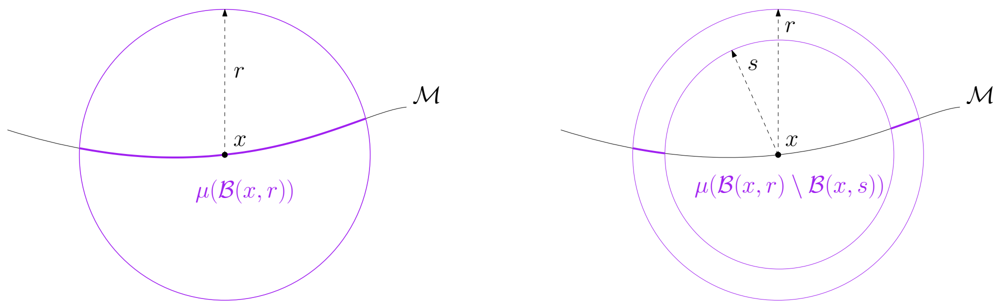

Figure 17: Representation of Proposition 2.17 first point (left) and third point (right).

Proof. Consider the map $\overline{\exp}_x^{\mathcal{M}}$ and the measure ν_x as defined in Lemma 2.15. In the following, we write $T = T_x\mathcal{M}$, and $\overline{\mathcal{B}}^T = \left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}\left(\overline{\mathcal{B}}(x, r)\right)$.

Point (1): We have $\mu\left(\overline{\mathcal{B}}(x, r)\right) = \nu_x\left(\overline{\mathcal{B}}^T\right)$. Writing down the density g of ν_x yields

$$\nu_x\left(\overline{\mathcal{B}}^T\right) = \int_{\overline{\mathcal{B}}^T} g(v) \mathrm{d}\mathcal{H}^d(v).$$

According to the expression of g in Lemma 2.15, we have $g \geq f_{\min} J_{\min}$. Therefore,

$$\int_{\overline{\mathcal{B}}^T} g(v) \mathrm{d}\mathcal{H}^d(v) \geq \int_{\overline{\mathcal{B}}^T} f_{\min} J_{\min} \mathrm{d}\mathcal{H}^d(v) = f_{\min} J_{\min} \mathcal{H}^d\left(\overline{\mathcal{B}}^T\right).$$

Besides, since $\overline{\mathcal{B}}^T \supset \overline{\mathcal{B}}_T(0, r)$, we have

$$\mathcal{H}^d\left(\overline{\mathcal{B}}^T\right) \geq \mathcal{H}^d\left(\overline{\mathcal{B}}_T(0, r)\right) = V_d r^d.$$

We finally obtain $\nu_x\left(\overline{\mathcal{B}}^T\right) \geq f_{\min} J_{\min} V_d r^d$.

Point (2): Observe that $\int_{\overline{\mathcal{B}}_T(0,r)} f(x) \mathrm{d}\mathcal{H}^d(v) = f(x) V_d r^d$. Hence

$$\begin{aligned} \left|\mu(\overline{\mathcal{B}}(x, r)) - f(x) V_d r^d\right| &= \left|\int_{\overline{\mathcal{B}}^T} g(v) \mathrm{d}\mathcal{H}^d(v) - \int_{\overline{\mathcal{B}}_T(0,r)} f(x) \mathrm{d}\mathcal{H}^d(v)\right| \\ &\leq \underbrace{\left|\int_{\overline{\mathcal{B}}_T(0,r)} (f(x) - g(v)) \mathrm{d}\mathcal{H}^d(v)\right|}_{(1)} + \underbrace{\left|\int_{\overline{\mathcal{B}}^T \setminus \overline{\mathcal{B}}_T(0,r)} g(v) \mathrm{d}\mathcal{H}^d(v)\right|}_{(2)}. \end{aligned}$$

To bound Term (1), notice that $g(0) = f(\exp_x^{\mathcal{M}}(0))J_0 = f(x)$. Hence we can write:

$$\left| \int_{\overline{\mathcal{B}}_T(0,r)} (f(x) - g(v)) \mathrm{d}\mathcal{H}^d(v) \right| \leq \int_{\overline{\mathcal{B}}_T(0,r)} \left| g(0) - g(v) \right| \mathrm{d}\mathcal{H}^d(v).$$

Now, Lemma 2.15 gives $|g(v) - g(0)| \leq c_{2.15} r$, and we obtain $\left| \int_{\overline{\mathcal{B}}_T(0,r)} (f(x) - g(v)) \mathrm{d}\mathcal{H}^d(v) \right| \leq c_{2.15} r V_d r^d$.

On the other hand, we bound Term (2) thanks to the inclusion $\overline{\mathcal{B}}^T \subseteq \overline{\mathcal{B}}_T\left(0, c_{2.10}(\rho r) r\right)$. Denote $\mathcal{A} = \overline{\mathcal{B}}_T\left(0, c_{2.10}(\rho r) r\right) \setminus \overline{\mathcal{B}}_T\left(0, r\right)$. We have $\overline{\mathcal{B}}^T \setminus \overline{\mathcal{B}}_T\left(0, r\right) \subset \mathcal{A}$, hence

$$\begin{aligned} \int_{\overline{\mathcal{B}}^T \setminus \overline{\mathcal{B}}_T(0,r)} g(v) \mathrm{d}\mathcal{H}^d(v) &\leq \int_{\mathcal{A}} g(v) \mathrm{d}\mathcal{H}^d(v) \\ &\leq f_{\max} J_{\max} \mathcal{H}^d(\mathcal{A}). \end{aligned}$$

Moreover, we have

$$\begin{aligned} \mathcal{H}^d(\mathcal{A}) &= \mathcal{H}^d\left(\overline{\mathcal{B}}_T\left(0, c_{2.10}(\rho r) r\right)\right) - \mathcal{H}^d\left(\overline{\mathcal{B}}_T\left(0, r\right)\right) \\ &= V_d \left(c_{2.10}(\rho r)^d - 1\right) r^d. \end{aligned}$$

We can use $c_{2.10}(\rho r) \leq 1 + 2\rho r \leq 2$ and the inequality $A^d - 1 \leq d(A-1)A^{d-1}$, where $A \geq 1$, to get

$$\begin{aligned} \left(c_{2.10}(\rho r)^d - 1\right) &\leq d \cdot \left(c_{2.10}(\rho r) - 1\right) \cdot c_{2.10}(\rho r)^{d-1} \\ &\leq d \cdot 2\rho r \cdot 2^{d-1}. \end{aligned}$$

We finally deduce the following bound on Term (2):

$$\int_{\overline{\mathcal{B}}^T \setminus \overline{\mathcal{B}}_T(0,r)} g(v) \mathrm{d}\mathcal{H}^d(v) \leq f_{\max} J_{\max} V_d r^d d \cdot \rho r 2^d.$$

Gathering Term (1) and (2), we obtain

$$\left| \mu(\overline{\mathcal{B}}\left(x, r\right)) - f(x) V_d r^d \right| \leq r \left(c_{2.15} + f_{\max} J_{\max} d \rho 2^d\right) V_d r^d.$$

Point (3): Let us write

$$\begin{aligned} \mu\left(\overline{\mathcal{B}}\left(x, r\right) \setminus \overline{\mathcal{B}}\left(x, s\right)\right) &= \nu_x\left(\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}\left(\overline{\mathcal{B}}\left(x, r\right) \setminus \overline{\mathcal{B}}\left(x, s\right)\right)\right) \\ &= \int_{\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}\left(\overline{\mathcal{B}}(x,r) \setminus \overline{\mathcal{B}}(x,s)\right)} g(v) \mathrm{d}\mathcal{H}^d(v). \end{aligned}$$

In spherical coordinates, this integral reads

$$\int_{\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}\left(\overline{\mathcal{B}}(x,r) \setminus \overline{\mathcal{B}}(x,s)\right)} g(v) \mathrm{d}\mathcal{H}^d(v) = \int_{v \in \partial \mathcal{B}_T(0,1)} \int_{t=a(v)}^{b(v)} g(tv) t^{d-1} \mathrm{d}t \mathrm{d}v, \tag{5}$$

where a and b are defined as follows: for every $v \in \partial \mathcal{B}_T\left(0, 1\right) \subset T_x\mathcal{M}$, let γ_0 be a arc-length parametrized geodesic with $\gamma(0) = x$ and $\dot{\gamma}(0) = v$, and set $a(v)$ and $b(v)$ to be the first positive values such that $\|\gamma(a(v)) - x\| = s$ and $\|\gamma(b(v)) - x\| = r$.

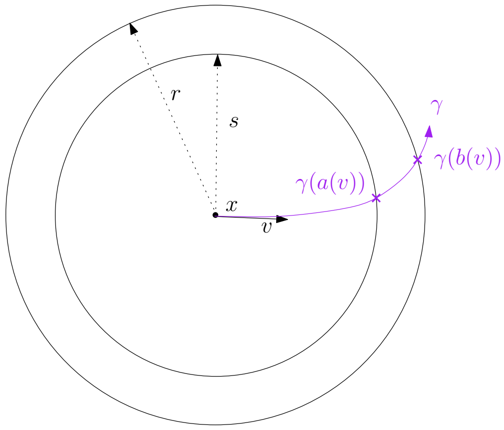

Figure 18: Illustration of $a(v)$ and $b(v)$ in the proof of Proposition 2.17.

Let us show that

$$b(v) - a(v) \leq \frac{1}{1-\rho r}(r-s) \tag{6}$$

Consider the application $\phi\colon t \mapsto \|\gamma(t)-x\|^2$. According to Lemma 2.6 Point 3 with $l = 0$, we have $\ddot{\phi}(t) \geq 2(1-\rho r)$ for $t \in [0, b(v)]$. It follows that $\dot{\phi}(t) \geq 2(1-\rho r)t$, and that

$$\begin{aligned}\phi(b(v)) - \phi(a(v)) = \int_{a(v)}^{b(v)} \dot{\phi}(t)\mathrm{d}t &\geq \int_{a(v)}^{b(v)} 2(1-\rho r)t\mathrm{d}t \\ &= (1-\rho r)(b(v)^2 - a(v)^2).\end{aligned}$$

Since $r^2 - s^2 = \phi(b(v)) - \phi(a(v))$, we deduce that

$$r^2 - s^2 \geq (1-\rho r)(b(v)^2 - a(v)^2). \tag{7}$$

Writing $r^2 - s^2 = (r+s)(r-s)$ and $b(v)^2 - a(v)^2 = \big(b(v)+a(v)\big)\big(b(v)-a(v)\big)$ leads to

$$(r-s)\frac{1}{1-\rho r}\frac{r+s}{b(v)+a(v)} \geq b(v) - a(v).$$

But $b(v) + a(v) \geq r + s$, hence $(r-s)\frac{1}{1-\rho r} \geq b(v) - a(v)$, as wanted.

Now, notice that we have $b(v) \leq 2r$. Indeed, $b < \frac{1}{\rho}$ by Lemma 2.6 Point 5 with $l = 0$, and we conclude with Lemma 2.3 Point 2. Hence we have

$$\int_{t=a(v)}^{b(v)} g(tv)t^{d-1}\mathrm{d}t \leq \int_{t=a(v)}^{b(v)} f_{\max} J_{\max}(2r)^{d-1}\mathrm{d}t.$$

Using Equation 6, we get

$$\begin{aligned}\int_{t=a(v)}^{b(v)} f_{\max} J_{\max}(2r)^{d-1}\mathrm{d}t &= (b(v)-a(v)) f_{\max} J_{\max}(2r)^{d-1}\mathrm{d}t \\ &\leq \frac{1}{1-\rho r}(r-s) f_{\max} J_{\max}(2r)^{d-1}.\end{aligned}$$

From these last two equations we deduce

$$\begin{aligned}\int_{v\in\partial\mathcal{B}(0,1)} \int_{t=a(v)}^{b(v)} g(tv)t^{d-1}\mathrm{d}t\mathrm{d}v &\leq \frac{1}{1-\rho r}(r-s) f_{\max} J_{\max}(2r)^{d-1} \int_{v\in\partial\mathcal{B}(0,1)} \mathrm{d}v \\ &= \frac{1}{1-\rho r}(r-s) f_{\max} J_{\max}(2r)^{d-1} \cdot dV_d.\end{aligned}$$

Going back to Equation 5, we obtain

$$\mu\left(\overline{\mathcal{B}}(x,r)\setminus\overline{\mathcal{B}}(x,s)\right)=\frac{2^{d-1}dV_d f_{\max}J_{\max}}{1-\rho r}(r-s)r^{d-1},$$

and we conclude with $r\leq\frac{1}{2\rho}$:

$$\mu\left(\overline{\mathcal{B}}(x,r)\setminus\overline{\mathcal{B}}(x,s)\right)=2^d dV_d f_{\max}J_{\max}(r-s)r^{d-1}.\qquad\square$$

The following proposition is a weaker form of Proposition 2.17, without normal reach condition. Its proof, based on the same ideas, is given in Appendix B.

Proposition 2.18. *Let $x_0\in\mathcal{M}_0$, $r\leq\frac{1}{2\rho}$ and $s\in[0,r]$. We have*

- $\mu\left(\overline{\mathcal{B}}(x,r)\right)\geq c_5 r^d$
- $\mu\left(\overline{\mathcal{B}}(x,r)\setminus\overline{\mathcal{B}}(x,s)\right)\leq c_7 r^{d-\frac{1}{2}}(r-s)^{\frac{1}{2}}$

with $c_5=f_{\min}J_{\min}V_d$ and $c_7=\frac{f_{\max}J_{\max}}{f_{\min}J_{\min}}(\frac{\rho}{\sqrt{4-\sqrt{13}}})^d d2^{2d}\sqrt{3}$.

2.5 Quantification of the normal reach

In this subsection, we suppose that the dimension of the manifold $\mathcal{M}_0$ is $d=1$, and we assume the Hypotheses 1, 2 and 3. We give an upper bound on the measure $\mu_0(\lambda_0^t)$, i.e., the measure of points $x_0\in\mathcal{M}_0$ with normal reach not greater than t. This proves a result announced in Subsection 1.2: Hypothesis 4 is a consequence of Hypotheses 1, 2 and 3.

We shall use two quantities related to the immersion $\mathcal{M}_0$. Let $\mathcal{D}_0$ be the set of critical points of the Euclidean distance on $\mathcal{M}_0$, that is,

$$\mathcal{D}_0=\left\{(x_0,y_0)\in\mathcal{M}_0,\ x_0\neq y_0,\ x-y\perp T_y\mathcal{M}\ \text{ and }\ x-y\perp T_x\mathcal{M}\right\}.\tag{8}$$

Also, let $\mathcal{C}_0$ be the set of self-intersections of $\mathcal{M}_0$:

$$\mathcal{C}_0=\left\{(x_0,y_0)\in\mathcal{M}_0,\ x_0\neq y_0\ \text{ and }\ x=y\right\}.\tag{9}$$

As a consequence of Remark 2.4 and the compacity of $\mathcal{M}_0$, the set $\mathcal{C}_0$ is finite. For every $(x_0,y_0)\in\mathcal{C}_0$, let $\theta(x_0,y_0)\in\left[0,\frac{\pi}{2}\right]$ be the angle formed by the lines $T_x\mathcal{M}$ and $T_y\mathcal{M}$. Define

$$\Theta=\inf\left\{\theta(x_0,y_0),\ (x_0,y_0)\in\mathcal{C}_0\right\}.\tag{10}$$

Note that, according to Hypothesis 1, we have $\Theta>0$. Besides, on the set $\mathcal{D}_0\setminus\mathcal{C}_0$, consider the quantity

$$\Delta=\inf\left\{\|x-y\|,\ (x_0,y_0)\in\mathcal{D}_0\setminus\mathcal{C}_0\right\}.\tag{11}$$

Since $\mathcal{C}_0$ consists of isolated points of $\mathcal{D}_0$, the set $\mathcal{D}_0\setminus\mathcal{C}_0$ is closed, hence the previous infimum is attained. Therefore, $\Delta>0$.

In order to bound the measure $\mu_0\left(\lambda_0^t\right)$, we first prove that the sublevel set λ_0^t is included in a thickening of $\mathcal{C}_0$ (Lemma 2.21). By bounding the measure of this thickening, we obtain the main result of this subsection (Proposition 2.22). We start by a lemma which describes the situation around self-intersection points of $\mathcal{M}_0$.

Lemma 2.19. *Let $(x_0^*,y_0^*)\in\mathcal{C}_0$. Denote by θ the angle formed by the lines $T_{x^*}\mathcal{M}$ and $T_{y^*}\mathcal{M}$. Let $x_0,y_0\in\mathcal{M}_0$. Denote $\delta=\mathrm{d}_{\mathcal{M}_0}(x_0^*,x_0)$ and $\delta'=\mathrm{d}_{\mathcal{M}_0}(y_0^*,y_0)$. If $\delta'\leq\delta\leq\frac{\sin(\theta)}{2\rho}$, then $\|x-y\|\geq\frac{\sin(\theta)}{2}\delta$.*

Proof. Let γ_0 be an arc-length parametrized geodesic connecting x_0^* to x_0, and η_0 connecting y_0^* to y_0. Let $v_0 = \dot{\gamma}_0(0)$, and $\overline{x} = x^* + \delta v$. Accordingly, let $w_0 = \dot{\eta}_0(0)$, and $\overline{y} = y^* + \delta' w = x^* + \delta' w$.

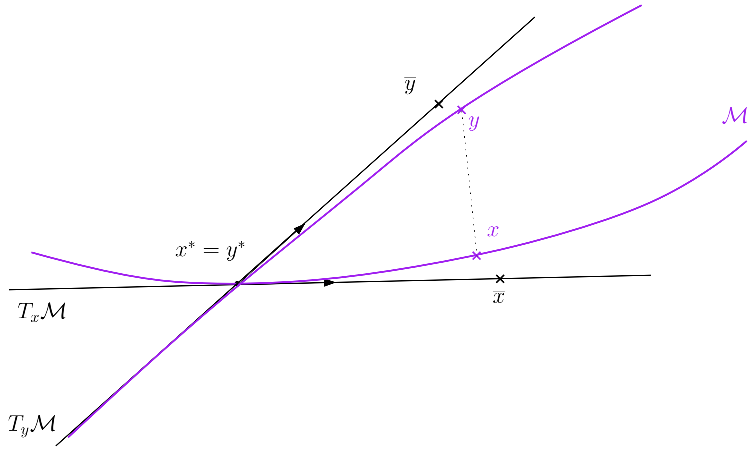

Figure 19: Situation in Lemma 2.19.

The triangular inequality yields

$$\|x - y\| \geq \|\overline{x} - \overline{y}\| - \|x - \overline{x}\| - \|y - \overline{y}\| .$$

According to Lemma 2.3, we have $\|x - \overline{x}\| \leq \frac{\rho}{2}\delta^2$ and $\|y - \overline{y}\| \leq \frac{\rho}{2}\delta'^2 \leq \frac{\rho}{2}\delta^2$. Moreover, $\|\overline{x} - \overline{y}\|$ is not lower than $\|\overline{x} - z\|$, where z is the projection of $\overline{x}$ on the line $T_{y^*}\mathcal{M}$. Elementary trigonometry shows that $\|\overline{x} - z\| = \sin(\theta)\delta$. Hence the previous Equation yields

$$\begin{aligned} \|x - y\| &\geq \sin(\theta)\delta - \frac{\rho}{2}\delta^2 - \frac{\rho}{2}\delta^2 \\ &= \sin(\theta)\delta \left(1 - \frac{\rho}{\sin(\theta)}\delta\right), \end{aligned}$$

and we conclude using $\delta \leq \frac{\sin(\theta)}{2\rho}$. □

Remark 2.20. A similar proof leads the following result: let $x_0, y_0, z_0 \in \mathcal{M}_0$. Denote $\delta = \mathrm{d}_{\mathcal{M}_0}(x_0^*, x_0)$ and $\delta' = \mathrm{d}_{\mathcal{M}_0}(y_0^*, y_0)$. Suppose that x_0 and y_0 are in opposite orientation around z_0, that is, there exist a unit vector $v \in T_{z_0}\mathcal{M}_0$ such that $x_0 = \exp_{z_0}^{\mathcal{M}_0}(\delta v)$ and $y_0 = \exp_{z_0}^{\mathcal{M}_0}(-\delta' v)$. If $\delta', \delta \leq \frac{1}{\rho}$, then $\|x - y\| \geq \frac{1}{2}(\delta + \delta')$.

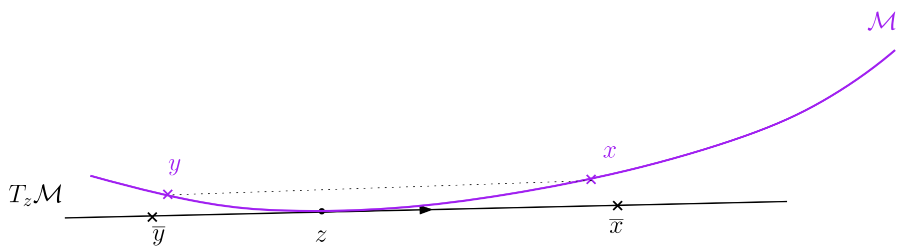

Figure 20: Situation in Remark 2.20.

The following lemma associates every point of $\mathcal{M}_0$ with small normal reach to a point with zero normal reach.

Lemma 2.21. *Let $x_0 \in \mathcal{M}_0$ with $\lambda_0(x_0) < \Delta \wedge \frac{\sin(\Theta)^2}{4\rho}$. Then there exists $x_0^* \in \mathcal{M}_0$ with*

$$\lambda_0(x_0^*) = 0 \qquad \textit{and} \qquad \mathrm{d}_{\mathcal{M}_0}(x_0, x_0^*) \leq c_{2.21}\lambda_0(x_0),$$

where $c_{2.21} = \frac{2}{\sin(\Theta)}$.

Proof. Let $y_0 \in \mathcal{M}_0$ such that $\|x - y\| = \lambda_0(x_0)$ and $x - y \perp T_y\mathcal{M}$. In order to find a point x_0^*, consider the following vector field on $\mathcal{M}_0 \times \mathcal{M}_0$:

$$\begin{array}{ccc} \mathcal{M}_0 \times \mathcal{M}_0 & \longrightarrow & T\mathcal{M}_0 \times T\mathcal{M}_0 \\ \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} & \longmapsto & \begin{pmatrix} p_{T_x\mathcal{M}}(y - x) \\ p_{T_y\mathcal{M}}(x - y) \end{pmatrix}, \end{array}$$

where $p_{T_x\mathcal{M}}$ and $p_{T_y\mathcal{M}}$ denote the orthogonal projection on $T_x\mathcal{M}$ and $T_y\mathcal{M}$. We implicitely use the identifications $T_x\mathcal{M} \simeq T_{x_0}\mathcal{M}_0$. Since $\mathcal{M}_0$ is $\mathcal{C}^2$, this vector field is of regularity $\mathcal{C}^1$, and we can apply Cauchy-Lipschitz theorem. Let u_0 be a maximal integral curve for this field, with initial value $u_0(0) = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$. Since $\mathcal{M}_0 \times \mathcal{M}_0$ is compact, the solution u_0 is global.

In order to study the convergence of u_0, we shall consider a Lyapunov map. Let $H \colon E \to \mathbb{R}$ be defined as $H(u) = \|u\|^2$. A computation shows that

$$\begin{aligned} H(\gamma(t) - \eta(t))' &= -2 \left\langle\, p_{T_{\gamma(t)}\mathcal{M}}(\gamma(t) - \eta(t)) \;,\; \gamma(t) - \eta(t) \,\right\rangle \\ &\quad - 2 \left\langle\, p_{T_{\eta(t)}\mathcal{M}}(\gamma(t) - \eta(t)) \;,\; \gamma(t) - \eta(t) \,\right\rangle \\ &= -2 \left\| p_{T_{\gamma(t)}\mathcal{M}}(\gamma(t) - \eta(t)) \right\|^2 - 2 \left\| p_{T_{\eta(t)}\mathcal{M}}(\gamma(t) - \eta(t)) \right\|^2. \end{aligned} \tag{12}$$

This quantity is nonpositive, hence the map $t \mapsto H(\gamma(t) - \eta(t))$ is nonincreasing. Note that for $t = 0$, we have $H(\gamma(0) - \eta(0)) = \lambda_0(x_0)$. Note also that for every $t \in \mathbb{R}^+$, we have $H(\gamma(t) - \eta(t)) \neq 0$, since the relation $\gamma(t) = \eta(t)$ corresponds to a stationary point of the system.

We divide the rest of the proof in five steps.

Step 1. Let us prove that $\mathrm{d}_{\mathcal{M}_0}(\gamma_0(t), \eta_0(t)) > \frac{1}{\rho}$ for every $t \in \mathbb{R}^+$. By contradiction, suppose that $\mathrm{d}_{\mathcal{M}_0}(\gamma_0(t), \eta_0(t)) \leq \frac{1}{\rho}$ for some t. As a consequence of Remark 2.5, we have $\mathrm{d}_{\mathcal{M}_0}(\gamma_0(0), \eta_0(0)) \geq \frac{1}{\rho}$. Therefore there exists a value $s \in [0, t]$ such that $\mathrm{d}_{\mathcal{M}_0}(\gamma_0(s), \eta_0(s)) = \frac{1}{\rho}$.

Let z_0 be a (geodesic) midpoint between $\gamma_0(s)$ and $\eta_0(s)$. We have

$$\mathrm{d}_{\mathcal{M}_0}(\gamma_0(s), z_0) = \mathrm{d}_{\mathcal{M}_0}(\eta_0(s), z_0) = \frac{1}{2\rho},$$

hence we can apply Remark 2.20 to get

$$\|\gamma(s) - \eta(s)\| \geq \frac{1}{2}\left(\mathrm{d}_{\mathcal{M}_0}(\gamma_0(s), z_0) + \mathrm{d}_{\mathcal{M}_0}(\gamma_0(s), z_0)\right) = \frac{1}{2\rho}.$$

Besides, we have seen that the map $t \mapsto \|\gamma(t) - \eta(t)\|$ is bounded above by $\|\gamma(0) - \eta(0)\| = \lambda_0(x_0)$. The inequality $\frac{1}{2\rho} \leq \|\gamma(s) - \eta(s)\| \leq \lambda_0(x_0)$ now contradicts the assumption $\lambda_0(x_0) < \frac{\sin(\Theta)^2}{4\rho}$.

Step 2. Let us show that $\gamma(t) - \eta(t)$ goes to zero. Let v_0 denote the map $v_0(t) = \gamma_0(t) - \eta_0(t)$, and $v(t) = \gamma(t) - \eta(t)$. It is enough to show that H is a strict Lyapunov map, i.e., there exists a constant $c > 0$ such that

$$H(v(t))' \leq -cH(v(t)). \tag{13}$$

According to Equation 12, we can write $H(v(t))' = -2c(t)\|v(t)\|^2$ with

$$c(t) = \frac{1}{\|v(t)\|^2}\left(\left\|p_{T_{\gamma(t)}\mathcal{M}}(v(t))\right\|^2 + \left\|p_{T_{\eta(t)}\mathcal{M}}(v(t))\right\|^2\right) \tag{14}$$

$$= \left\|p_{T_{\gamma(t)}\mathcal{M}}\left(\frac{v(t)}{\|v(t)\|}\right)\right\|^2 + \left\|p_{T_{\eta(t)}\mathcal{M}}\left(\frac{v(t)}{\|v(t)\|}\right)\right\|^2. \tag{15}$$

To prove Equation 13, it remains to show that $c(t)$ is bounded below.

By contradiction, suppose that it is not the case. This implies that there exists an increasing sequence $(t_n)_{n\geq 0}$ such that the sequence $(c(t_n))_{n\geq 0}$ goes to 0. By compacity of $\mathcal{M}_0$, we can assume that $(x_0(t_n))_{n\geq 0}$ and $(y_0(t_n))_{n\geq 0}$ admit a limit, that we denote x_0^* and y_0^*. By compacity of the unit sphere of E, we can also assume that $\left(\frac{v(t_n)}{\|v(t_n)\|}\right)_{n\geq 0}$ admits a limit v^*, as well as $\left(\frac{\gamma(t_n)}{\|v(t_n)\|}\right)_{n\geq 0}$ and $\left(\frac{\eta(t_n)}{\|v(t_n)\|}\right)_{n\geq 0}$. Note already the following facts: $\|v^*\| = 1$, and v^* is included in the 2-dimensional affine space spanned by $T_{x^*}\mathcal{M}$ and $T_{y^*}\mathcal{M}$.

According to Step 1, we have $x_0^* \neq y_0^*$. Let us prove that $x^* = y^*$. By contradiction suppose that it is not the case. Then $(v(t_n))_{n\geq 0}$ goes to the nonzero vector $x^* - y^*$. Using that $c(t_n)$ goes to zero, Equation 14 yields

$$\|p_{T_{x^*}\mathcal{M}}(x^* - y^*)\| = \left\|p_{T_{y^*}\mathcal{M}}(x^* - y^*)\right\| = 0.$$

Hence the pair (x^*, y^*) is an element of $\mathcal{D}_0$ (defined in Equation 8). By definition of Δ (Equation 11), we obtain $\|x^* - y^*\| \geq \Delta$. Besides, since the map $t \mapsto \|\gamma(t) - \eta(t)\|$ is non-increasing, we get $\|x^* - y^*\| \leq \|x - y\|$, which is lower than Δ by assumption. This is a contradiction.

Now, we have $x^* = y^*$. Still using that $c(t_n)$ goes to zero, Equation 15 yields

$$\|p_{T_{x^*}\mathcal{M}}(v^*)\| = \left\|p_{T_{y^*}\mathcal{M}}(v^*)\right\| = 0.$$

But $x^* = y^*$ implies that $T_{x^*}\mathcal{M} \neq T_{y^*}\mathcal{M}$, according to Hypothesis 1. In conclusion, v^* is a vector of the affine space spanned by $T_{x^*}\mathcal{M}$ and $T_{y^*}\mathcal{M}$, and v^* is orthogonal to both these lines. Hence v^* has to be zero, which is absurd since it has norm 1. We deduce that $c(t)$ is bounded below, and that H is a strict Lyapunov map.

Step 3. Let us prove that u_0 admits a limit $\begin{pmatrix} x_0^* \\ y_0^* \end{pmatrix}$ when $t \to +\infty$, with $x_0^* \neq y_0^*$ and $x^* = y^*$. By compacity of $\mathcal{M}_0 \times \mathcal{M}_0$, we can pick two accumulation points x_0^* and y_0^* of γ_0 and η_0. Let us prove that, for every $\epsilon > 0$, there exists a $t \geq 0$ such that for every $s \geq t$, the geodesic distances $\mathrm{d}_{\mathcal{M}_0}(\gamma_0(s), x_0^*)$ and $\mathrm{d}_{\mathcal{M}_0}(\eta_0(s), y_0^*)$ are upper bounded by ϵ. This would imply that γ_0 and η_0 admit x_0^* and y_0^* as limits. Let $\epsilon > 0$. We can assume that $\epsilon < \frac{\sin(\Theta)}{2\rho}$, where Θ is defined in Equation 10.

According to Step 2, we have $x^* = y^*$. Hence the tangent spaces $T_{x^*}\mathcal{M}$ and $T_{y^*}\mathcal{M}$ are different. Let $\theta \in \left(0, \frac{\pi}{2}\right]$ be the angle they form. Since the map $t \mapsto \|\gamma(t) - \eta(t)\|$ goes to zero, there exists a $t \geq 0$ such that for every $s \geq t$, we have

$$\|\gamma(t) - \eta(t)\| < \frac{\sin(\theta)}{2}\epsilon. \tag{16}$$

Now, by definition of the accumulation points x^* and y^*, there exists a $t' \geq t$ such that

$$\mathrm{d}_{\mathcal{M}_0}(\gamma_0(t'), x_0^*) \leq \epsilon \qquad \text{and} \qquad \mathrm{d}_{\mathcal{M}_0}(\eta_0(t'), y_0^*) \leq \epsilon. \tag{17}$$

We shall deduce that for every $s \geq t'$, we have

$$\mathrm{d}_{\mathcal{M}_0}(\gamma_0(s), x_0^*) \leq \epsilon \qquad \text{and} \qquad \mathrm{d}_{\mathcal{M}_0}(\eta_0(s), y_0^*) \leq \epsilon. \tag{18}$$

Let us prove it by contradiction. From Equation 17 and the assumption that Equation 18 is false, we deduce that there exist a first value $s \geq t'$ such that $\delta = \mathrm{d}_{\mathcal{M}_0}(\gamma_0(s), x_0^*) = \epsilon$ or $\delta' = \mathrm{d}_{\mathcal{M}_0}(\eta_0(s), x_0^*) = \epsilon$. Since $\epsilon < \frac{\sin(\Theta)}{2\rho}$, we can use Lemma 2.19 to deduce

$$\|\gamma_0(s) - \eta_0(s)\| \geq \frac{\sin(\theta)}{2}\epsilon.$$

But this contradicts Equation 16.

Step 4. Let us show that $\mathrm{d}_{\mathcal{M}_0}(x_0, x_0^*) \leq \frac{\sin(\theta)}{2\rho}$ and $\mathrm{d}_{\mathcal{M}_0}(y_0, y_0^*) \leq \frac{\sin(\theta)}{2\rho}$. By contradiction, suppose that $\mathrm{d}_{\mathcal{M}_0}(x_0, x_0^*) > \frac{\sin(\theta)}{2\rho}$ or $\mathrm{d}_{\mathcal{M}_0}(y_0, y_0^*) > \frac{\sin(\theta)}{2\rho}$. According to the limits $\gamma_0 \to x_0^*$ and $\eta_0 \to y_0^*$, there exists $t \in \mathbb{R}^+$ such that

$$\begin{aligned} & \mathrm{d}_{\mathcal{M}_0}(\gamma_0(t), x_0^*) = \frac{\sin(\theta)}{2\rho} \quad \text{and} \quad \mathrm{d}_{\mathcal{M}_0}(\eta_0(t), y_0^*) \leq \frac{\sin(\theta)}{2\rho} \\ \text{or} \quad & \mathrm{d}_{\mathcal{M}_0}(\gamma_0(t), x_0^*) \leq \frac{\sin(\theta)}{2\rho} \quad \text{and} \quad \mathrm{d}_{\mathcal{M}_0}(\eta_0(t), y_0^*) = \frac{\sin(\theta)}{2\rho}. \end{aligned}$$

In both cases, we can apply Lemma 2.19 to get

$$\|\gamma(t) - \eta(t)\| \geq \frac{\sin(\theta)}{2} \cdot \frac{\sin(\theta)}{2\rho} = \frac{\sin(\theta)^2}{4\rho}. \tag{19}$$

Since the map $t \mapsto \|\gamma(t) - \eta(t)\|$ is non-increasing, we have

$$\|\gamma(t) - \eta(t)\| \leq \|\gamma(0) - \eta(0)\| = \|x - y\| = \lambda_0(x_0).$$

But $\lambda_0(x_0) < \frac{\sin(\Theta)^2}{4\rho}$ by assumption. Hence $\|\gamma(t) - \eta(t)\| < \frac{\sin(\theta)^2}{4\rho}$, which contradicts Equation 19.

Step 5. Let us show that $\mathrm{d}_{\mathcal{M}_0}(x_0, x_0^*) \geq \frac{2}{\sin(\theta)}\lambda_0(x_0)$. According to Step 4, we have $\mathrm{d}_{\mathcal{M}_0}(x_0, x_0^*) \leq \frac{\sin(\theta)}{2\rho}$ and $\mathrm{d}_{\mathcal{M}_0}(y_0, y_0^*) \leq \frac{\sin(\theta)}{2\rho}$. Therefore, Lemma 2.19 gives

$$\|x - y\| \geq \frac{\sin(\theta)}{2}\mathrm{d}_{\mathcal{M}_0}(x_0, x_0^*).$$

Using $\|x - y\| = \lambda_0(x_0)$ and $\sin(\theta) \geq \sin(\Theta)$, we obtain

$$\mathrm{d}_{\mathcal{M}_0}(x_0, x_0^*) \leq \frac{2}{\sin(\theta)}\|x - y\| \leq \frac{2}{\sin(\Theta)}\lambda_0(x_0). \qquad \square$$

We can now deduce the main result of this subsection: Hypothesis 4 holds in dimension 1.

Proposition 2.22. *For every $r < \Delta \wedge \frac{\sin(\Theta)^2}{4\rho}$, we have*

$$\mu_0(\lambda_0^r) \leq c_{2.22} r$$

where $c_{2.22} = |C_0| f_{\max} J_{\max} c_{2.21}$ and $|C_0|$ is the number of self-intersection points of $\mathcal{M}_0$.

Proof. Let C_0 denote the set of self-intersection points of $\mathcal{M}_0$, i.e.,

$$C_0 = \{x_0 \in \mathcal{M}_0, \lambda_0(x_0) = 0\}.$$

Observe that C_0 is closely related to the set $\mathcal{C}_0$ defined in Equation 9. Using Lemma 2.21, we can pair every $x_0 \in \lambda_0^r$ to a point $x_0^* \in C_0$ with $\mathrm{d}_{\mathcal{M}_0}(x_0, x_0^*) \leq c_{2.21}\lambda_0(x_0)$. In other words, the sublevel set λ_0^r is included in the (geodesic) thickening

$$C_0^{c_{2.21}r} = \{x_0 \in \mathcal{M}_0, \exists x_0^* \in C_0, \mathrm{d}_{\mathcal{M}_0}(x_0, x_0^*) \leq r\}.$$

Now, C_0 is a finite set, and we write its thickening as a union of geodesic balls:

$$C_0^{c_{2.21}r} = \bigcup_{x_0 \in \mathcal{C}_0} \overline{\mathcal{B}}_{\mathcal{M}_0}(x_0, c_{2.21}r)$$

Thanks to Hypothesis 3, we can relate the measure μ_0 to the 1-dimensional Hausdorff measure $\mathcal{H}^1$. As in the proof of Proposition 2.17, we get

$$\mu_0\left(\overline{\mathcal{B}}_{\mathcal{M}_0}(x_0, c_{2.21}r)\right) \leq f_{\max} J_{\max} c_{2.21} r.$$

Therefore, if $|C_0|$ denotes the cardinal of C_0, we obtain

$$C_0^{c_{2.21}r} \leq |C_0| f_{\max} J_{\max} c_{2.21} r. \qquad \square$$

3 Tangent space estimation

In this section, we show that one can estimate the tangent spaces of $\mathcal{M}$ based on a sample of it, via the computation of local covariance matrices. We study the consistency of this estimation in Subsection 3.2, which is based on the results of the last section. In Subsection 3.3 we prove that this estimation is stable, based on lighter hypotheses than 1, 2 and 3.

3.1 Local covariance matrices and lifted measure

Definition 3.1. Let ν be any probability measure on E. Let $r > 0$ and $x \in \operatorname{supp}(\nu)$. The *local covariance matrix of ν around x at scale r* is the following matrix:

$$\Sigma_\nu(x) = \int_{\overline{\mathcal{B}}(x,r)} (x-y)^{\otimes 2} \frac{\mathrm{d}\nu(y)}{\nu(\overline{\mathcal{B}}(x,r))}$$

We also define the *normalized local covariance matrix* as $\overline{\Sigma}_\nu(x) = \frac{1}{r^2}\Sigma_\nu(x)$.

Note that $\Sigma_\nu(x)$ and $\overline{\Sigma}_\nu(x)$ depend on r, which is not made explicit in the notation. The normalization factor $\frac{1}{r^2}$ of the normalized local covariance matrix is justified by Proposition 3.1. Moreover, we introduce the following notations: for every $r > 0$ and $x \in \operatorname{supp}(\nu)$,

- ν_x is the restriction of ν to the ball $\overline{\mathcal{B}}(x,r)$,
- $\overline{\nu_x} = \frac{1}{\nu(\overline{\mathcal{B}}(x,r))}\nu_x$ is the corresponding probability measure.

Thus the local covariance matrix can be written as $\Sigma_\nu(x) = \int (x-y)^{\otimes 2}\mathrm{d}\overline{\nu_x}(y)$.

The collection of probability measures $\{\overline{\nu_x}\}_{x\in\operatorname{supp}(\nu)}$ is called in [MSW19, Section 3.3] the local truncation of ν at scale r. The application $x \mapsto \Sigma_\nu(x)$ is called in [MMM18, Section 2.2] the multiscale covariance tensor field of ν associated to the truncation kernel.

We remind the reader that the aim of this paper is to estimate the measure $\check{\mu}_0$, defined on $E\times\mathcal{M}(E)$ as $\check{\mu}_0 = \check{u}_*\mu_0$ (see Subsection 1.2). We call it the *exact lifted measure*. In other words, it can be defined as

$$\check{\mu}_0 = (u_*\mu_0)(x_0) \otimes \delta_{\frac{1}{d+2}p_{T_x\mathcal{M}}}$$

by disintegration of measure. Here is another alternative definition of $\check{\mu}_0$: for any $\phi\colon E\times\mathcal{M}(E) \to \mathbb{R}$ with compact support,

$$\int \phi(x,A)\mathrm{d}\check{\mu}_0(x,A) = \int \phi\left(u(x_0), \frac{1}{d+2}p_{T_x\mathcal{M}}\right)\mathrm{d}\mu_0(x_0). \tag{20}$$

In order to approximate $\check{\mu}_0$, we consider the following construction.

Definition 3.2. if ν is any measure on E, we denote by $\check{\nu}$ the measure on $E\times\mathcal{M}(E)$ defined by

$$\check{\nu} = \nu(x) \otimes \delta_{\overline{\Sigma}_\nu(x)}.$$

It is called the *lifted measure* associated to ν. In other words, for every $\phi\colon E\times\mathcal{M}(E)\to\mathbb{R}$ with compact support, we have

$$\int \phi(x,A)\mathrm{d}\check{\nu}(x,A) = \int \phi\Big(x, \overline{\Sigma}_\nu(x)\Big)\mathrm{d}\nu(x).$$

In accordance with the local covariance matrices, the lifted measure $\check{\nu}$ depends on the parameter r which is not made explicit in the notation. In order to compare these measures, we consider a Wasserstein-type distance on the space $E \times \mathcal{M}(E)$. Fix $\gamma > 0$, and let $\|\cdot\|_\gamma$ be the Euclidean norm on $E \times \mathcal{M}(E)$ defined as

$$\|(x, A)\|_\gamma^2 = \|x\|^2 + \gamma^2 \|A\|_{\mathrm{F}}^2 , \tag{21}$$

where $\|\cdot\|$ represents the usual Euclidean norm on E and $\|\cdot\|_{\mathrm{F}}$ represents the Frobenius norm on $\mathcal{M}(E)$. Let $p \geq 1$. We denote by $W_{p,\gamma}(\cdot,\cdot)$ the p-Wasserstein distance with respect to this metric. By definition, if α, β are probability measures on $E \times \mathcal{M}(E)$, then $W_{p,\gamma}(\alpha, \beta)$ can be written as

$$W_{p,\gamma}(\alpha, \beta) = \inf_\pi \left(\int_{(E\times\mathcal{M}(E))^2} \|(x, A) - (y, B)\|_\gamma^p \, \mathrm{d}\pi\big((x, A), (y, B)\big) \right)^{\frac{1}{p}} , \tag{22}$$

where the infimum is taken over all measures π on $(E \times \mathcal{M}(E))^2$ with marginals α and β.

We subdivise the rest of this section in three subsections. They respectively consists in showing that

- **Consistency:** if μ_0 is a measure satisfying the Hypotheses 2 and 3, then $W_{p,\gamma}(\check{\mu}_0, \check{\mu})$ is small (Proposition 3.4),
- **Stability:** in addition, if ν is a measure on E such that $W_p(\mu, \nu)$ is small, then so is $W_{p,\gamma}(\check{\mu}, \check{\nu})$ (Proposition 3.6)
- **Approximation:** under the previous hypotheses, $W_{p,\gamma}(\check{\mu}_0, \check{\nu})$ is small (Theorem 3.10).

The first point means that the lifted measure $\check{\mu}$ is close to the exact lifted measure $\check{\mu}_0$. In other words, construction we propose is consistent. If we are not observing μ but a close measure ν, the second point states that the lifted measure $\check{\nu}$ is still close to $\check{\mu}$. Combining these two statements gives the third one: the lifted measure $\check{\nu}$ is close the exact lifted measure $\check{\mu}_0$.

These several measures fit in a commutative diagram:

$$\begin{tikzcd}
\mathcal{M}_0 \arrow[rr, "\check{u}"] \arrow[dr, "u"'] & & E \times \mathcal{M}(E) \arrow[dl, "\mathrm{proj}"] \\
& E &
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
\mu_0 \arrow[rr, "\check{u}_*"] \arrow[dr, "u_*"'] & & \check{\mu}_0 & \check{\mu} & \check{\nu} \\
& \mu \arrow[ur, "g_*"] \arrow[urr, bend right, "(f_\mu)_*"'] & & & \nu \arrow[u, "(f_\nu)_*"']
\end{tikzcd}$$

where the maps g, f_μ and $f_\nu \colon E \to E \times \mathcal{M}(E)$ are defined as

$$g \colon x \longmapsto \left(x, \frac{1}{d+2} p_{T_x\mathcal{M}} \right), \qquad f_\mu \colon x \longmapsto \Big(x, \overline{\Sigma}_\mu(x) \Big), \qquad f_\nu \colon x \longmapsto \Big(x, \overline{\Sigma}_\nu(x) \Big).$$

Note that the map g is well-defined only on points $x \in \mathcal{M}$ that are not self-intersection points, i.e., points x such that $\lambda(x) > 0$. Under Hypothesis 4, g is well-defined μ-almost surely. The maps f_μ and f_ν are defined respectively on $\mathrm{supp}\,(\mu)$ and $\mathrm{supp}\,(\nu)$.

3.2 Consistency of the estimation

In this subsection, we assume that $\mathcal{M}_0$ and μ_0 satisfy the hypotheses 2 and 3.

The following proposition shows that the normalized covariance matrix approximates the tangent spaces of $\mathcal{M}$, as long as the parameter r is choosen smaller than the normal reach. A similar result appears in [ACLZ17, Lemma 13] in the case where $\mathcal{M}$ is a submanifold and μ is the uniform distribution on $\mathcal{M}$. Based on this result, we deduce that the lifted measure $\check{\mu}$ is close to the exact lifted measure $\check{\mu}_0$. The quality of this approximation depends on the measure of points with small normal reach, i.e., points where the tangent spaces are not well-estimated.

Proposition 3.1. *Let $x_0 \in \mathcal{M}_0$ and $r < \lambda(x) \wedge \frac{1}{2\rho}$. Denote by $p_{T_x\mathcal{M}}$ the orthogonal projection on the tangent space $T_x\mathcal{M}$, seen as a matrix. We have*

$$\left\| \overline{\Sigma}_\mu(x) - \frac{1}{d+2} p_{T_x\mathcal{M}} \right\|_{\mathrm{F}} \leq c_{3.3} r.$$

Proposition 3.1 is a direct consequence of the two following lemmas.

Lemma 3.2 ([ACLZ17, Lemma 11])**.** *The following matrix is equal to $r^2 \frac{1}{d+2} p_{T_x\mathcal{M}}$:*

$$\Sigma_* = \int_{\overline{\mathcal{B}}_{T_x\mathcal{M}}(0,r)} y^{\otimes 2} \frac{\mathrm{d}\mathcal{H}^d(y)}{V_d r^d}.$$

Lemma 3.3. *Still denoting $\Sigma_* = \int_{\overline{\mathcal{B}}_{T_x\mathcal{M}}(0,r)} y^{\otimes 2} \frac{\mathrm{d}\mathcal{H}^d(y)}{V_d r^d}$, we have*

$$\|\Sigma_\mu(x) - \Sigma_*\|_{\mathrm{F}} \leq c_{3.3} r^3,$$

where $c_{3.3} = 6\rho + 4\frac{c_{2.15}}{f_{\min} J_{\min}} + \frac{f_{\max}}{f_{\min} J_{\min}} 2^d d\rho + \frac{c_{2.17}}{f_{\min} J_{\min}}$.

Proof. We use the notations of Lemmas 2.15 and 2.17. We write $T = T_x\mathcal{M}$, $\overline{\mathcal{B}} = \overline{\mathcal{B}}(x,r)$ and $\overline{\mathcal{B}}^T = (\overline{\exp}_x^{\mathcal{M}})^{-1}(\overline{\mathcal{B}})$. We shall consider the following intermediate matrices:

$$\begin{aligned}
\Sigma_1 &= \int_{\overline{\mathcal{B}}} \left(\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x') \right)^{\otimes 2} \mathrm{d}\overline{\mu_x}(x') \\
\Sigma_2 &= \int_{\overline{\mathcal{B}}^T} g(0) \cdot y^{\otimes 2} \frac{\mathrm{d}\mathcal{H}^d(y)}{|\mu_x|} \\
\Sigma_3 &= \int_{\overline{\mathcal{B}}_T(0,r)} g(0) \cdot y^{\otimes 2} \frac{\mathrm{d}\mathcal{H}^d(y)}{|\mu_x|}
\end{aligned}$$

Let us write the triangle inequality:

$$\|\Sigma_\mu(x) - \Sigma_*\|_{\mathrm{F}} \leq \underbrace{\|\Sigma_\mu(x) - \Sigma_1\|_{\mathrm{F}}}_{(1)} + \underbrace{\|\Sigma_1 - \Sigma_2\|_{\mathrm{F}}}_{(2)} + \underbrace{\|\Sigma_2 - \Sigma_3\|_{\mathrm{F}}}_{(3)} + \underbrace{\|\Sigma_3 - \Sigma_*\|_{\mathrm{F}}}_{(4)}.$$

Term (1): By definition of the local covariance matrix, we have

$$\Sigma_\mu(x) = \int_{\overline{\mathcal{B}}(x,r)} (x - x')^{\otimes 2} \overline{\mu_x}(x').$$

We use the majoration

$$\begin{aligned}
\|\Sigma_\mu(x) - \Sigma_1\|_{\mathrm{F}} &\leq \int_{\overline{\mathcal{B}}(x,r)} \left\| (x-x')^{\otimes 2} - \left(\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x') \right)^{\otimes 2} \right\|_{\mathrm{F}} \mathrm{d}\overline{\mu_x}(x') \\
&\leq \sup_{x' \in \overline{\mathcal{B}}(x,r) \cap \mathcal{M}} \left\| (x-x')^{\otimes 2} - \left(\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x') \right)^{\otimes 2} \right\|_{\mathrm{F}}.
\end{aligned}$$

Let $x' \in \mathcal{B}(x,r) \cap \mathcal{M}$. According to Lemma 2.10, we have $\left\| \left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x') \right\| \leq 2r$. Moreover, $\|x - x'\| \leq r$, and Lemma C.1 gives

$$\left\| (x-x')^{\otimes 2} - \left(\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x') \right)^{\otimes 2} \right\|_{\mathrm{F}} \leq (r + 2r) \left\| (x' - x) - \left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x') \right\|. \tag{23}$$

Now, let us justify that

$$\left\| (x' - x) - \left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x') \right\| \leq \frac{\rho}{2} \mathrm{d}_{\mathcal{M}_0}(x_0, x_0')^2. \tag{24}$$

If we write $x' = \gamma(\delta)$ with γ a geodesic such that $\gamma(0) = x$ and $\delta = \mathrm{d}_{\mathcal{M}_0}(x_0, x_0')$, then $\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x') = \delta\dot{\gamma}(0)$, and we can write

$$\begin{aligned}\left\|(x'-x) - \left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x')\right\| &= \|\gamma(\delta) - (x + \delta\dot{\gamma}(0))\| \\ &\leq \frac{\rho}{2}\delta^2,\end{aligned}$$

where we used Lemma 2.3 for the last inequality. Hence Equation 24 is true. Combined with Lemma 2.10, which gives $\mathrm{d}_{\mathcal{M}_0}(x_0, x_0') \leq 2\|x - x'\| \leq 2r$, we obtain

$$\left\|(x-x')^{\otimes 2} - \left(\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x')\right)^{\otimes 2}\right\|_{\mathrm{F}} \leq \frac{\rho}{2}(2r)^2 = 2\rho r^2.$$

To conclude, we use Equation 23 to deduce $\|\Sigma_\mu(x) - \Sigma_1\|_{\mathrm{F}} \leq (r + 2r)2\rho r^2 = 6\rho r^3$.

Term (2): By transfert, we can write Σ_1 as

$$\Sigma_1 = \int_{\overline{\mathcal{B}}} \left(\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x')\right)^{\otimes 2} \frac{\mathrm{d}\mathcal{H}^d(y)}{|\mu_x|} = \int_{\overline{\mathcal{B}}^T} g(y) y^{\otimes 2} \frac{\mathrm{d}\mathcal{H}^d(y)}{|\mu_x|}.$$

We deduce the majoration

$$\|\Sigma_1 - \Sigma_2\|_{\mathrm{F}} \leq \int_{\overline{\mathcal{B}}^T} |g(0) - g(y)| \left\|y^{\otimes 2}\right\| \frac{\mathrm{d}\mathcal{H}^d(y)}{|\mu_x|}.$$

According to Lemma C.1, $\left\|y^{\otimes 2}\right\| = \|y\|^2 \leq (2r)^2$, and from Lemma 2.15 we get $|g(y) - g(0)| \leq c_{2.15} r$. Therefore,

$$\|\Sigma_1 - \Sigma_2\|_{\mathrm{F}} \leq 4r^2 \cdot c_{2.15} r \cdot \frac{\mathcal{H}^d\left(\overline{\mathcal{B}}^T\right)}{|\mu_x|}.$$

To conclude, note that $|\mu_x| \geq f_{\min} J_{\min} \mathcal{H}^d\left(\overline{\mathcal{B}}^T\right)$ (as in Lemma 2.15), so we obtain $\|\Sigma_1 - \Sigma_2\|_{\mathrm{F}} \leq 4\frac{c_{2.15}}{f_{\min} J_{\min}} r^3$.

Term (3): As for the previous terms, we use the majoration

$$\|\Sigma_2 - \Sigma_3\|_{\mathrm{F}} \leq \int_{\overline{\mathcal{B}}_T(0,r) \setminus \overline{\mathcal{B}}^T} \left\|g(0) \cdot y^{\otimes 2}\right\|_{\mathrm{F}} \frac{\mathrm{d}\mathcal{H}^d(y)}{|\mu_x|}.$$

One the one hand, $\left\|g(0) \cdot y^{\otimes 2}\right\|_{\mathrm{F}} \leq g(0) r^2 \leq f_{\max} r^2$, and we get

$$\|\Sigma_2 - \Sigma_3\|_{\mathrm{F}} \leq f_{\max} r^2 \frac{\mathcal{H}^d\left(\overline{\mathcal{B}}_T(0,r) \setminus \overline{\mathcal{B}}^T\right)}{|\mu_x|}.$$

On the other hand, since $\overline{\mathcal{B}}^T \subseteq \overline{\mathcal{B}}_T(x, c_{2.10}(\rho r) r)$, we have

$$\mathcal{H}^d\left(\overline{\mathcal{B}}^T \setminus \overline{\mathcal{B}}_T(0,r)\right) = (c_{2.10}(\rho r) r)^d V_d - r^d V_d.$$

The inequality $A^d - 1 \leq d(A-1)A^{d-1}$, where $A \geq 1$, gives

$$(c_{2.10}(\rho r) r)^d V_d - r^d V_d \leq V_d r^d \cdot d(c_{2.10}(\rho r) - 1) 2^{d-1}.$$

Combined with the inequalities $c_{2.10}(\rho r) \leq 1 + 2\rho r$ and $|\mu_x| \geq f_{\min} J_{\min} V_d r^d$, we get

$$\|\Sigma_2 - \Sigma_3\|_{\mathrm{F}} \leq \frac{f_{\max}}{f_{\min} J_{\min}} 2^d d \rho r^3.$$

Term (4): Let us write Σ^* as

$$\Sigma_* = \int_{\overline{\mathcal{B}}_{T_x\mathcal{M}}(0,r)} y^{\otimes 2} \frac{|\mu_x|}{V_d r^d} \frac{\mathrm{d}\mathcal{H}^d(y)}{|\mu_x|}.$$

Hence we have

$$\|\Sigma_3 - \Sigma_*\|_{\mathrm{F}} \leq \int_{\overline{\mathcal{B}}_T(0,r)} \left| \frac{|\mu_x|}{V_d r^d} - f(x) \right| \left\| y^{\otimes 2} \right\|_{\mathrm{F}} \frac{\mathrm{d}\mathcal{H}^d(y)}{|\mu_x|}.$$

According to Lemma 2.17 point 2, $\left| \frac{|\mu_x|}{V_d r^d} - f(x) \right| \leq c_{2.17} r$. Moreover, $\left\| y^{\otimes 2} \right\|_{\mathrm{F}} \leq r^2$ and $\int_{\overline{\mathcal{B}}_T(0,r)} \frac{\mathrm{d}\mathcal{H}^d(y)}{|\mu_x|} \leq \frac{1}{f_{\min} J_{\min}}$. Therefore,

$$\|\Sigma_3 - \Sigma_*\|_{\mathrm{F}} \leq \frac{c_{2.17}}{f_{\min} J_{\min}} r^3. \qquad \square$$

We now deduce a result concerning the lifted measures $\check{\mu}$ and $\check{\mu}_0$ (defined in Subsection 3.1). We remind the reader that the notation λ^r refers to the sublevel set $\lambda^{-1}([0,r])$. The quantity $\mu(\lambda^r)$ is the measure of points $x \in \mathcal{M}$ such that $\lambda(x) \leq t$.

Proposition 3.4. *Let $r < \frac{1}{2\rho}$. Then*

$$W_{p,\gamma}(\check{\mu}, \check{\mu}_0) \leq \gamma \left(2\mu(\lambda^r)^{\frac{1}{p}} + c_{3.1} r \right).$$

Proof. Define the map $\phi\colon \mathcal{M}_0 \to (E \times \mathcal{M}(E)) \times (E \times \mathcal{M}(E))$ as

$$\phi\colon x_0 \mapsto \left(\left(x, \overline{\Sigma}_\mu(x) \right), \left(x, \frac{1}{d+2} p_{T_x\mathcal{M}} \right) \right),$$

and consider the measure $\pi = \phi_* \mu_0$. It is a transport plan between $\check{\mu}$ and $\check{\mu}_0$. By definition of the Wasserstein distance, $W_{p,\gamma}^p(\check{\mu}, \check{\mu}_0) \leq \int \|(x,T) - (x',T')\|_\gamma^p \,\mathrm{d}\pi\left((x,T),(x',T')\right)$, and we can write

$$\begin{aligned} W_{p,\gamma}^p(\check{\mu}, \check{\mu}_0) &\leq \int \left\| \left(x, \frac{1}{r^2}\Sigma_\mu(x) \right) - \left(x, \frac{1}{d+2} p_{T_x\mathcal{M}} \right) \right\|_\gamma^p \mathrm{d}\mu(x) \\ &= \gamma^p \int \left\| \frac{1}{r^2}\Sigma_\mu(x) - \frac{1}{d+2} p_{T_x\mathcal{M}} \right\|_{\mathrm{F}}^p \mathrm{d}\mu(x). \end{aligned}$$

We split this last integral into the sets $A = \lambda^r$ and $B = E \setminus \lambda^r$.

On A, we use the majoration $\left\| \frac{1}{r^2}\Sigma_\mu(x) - \frac{1}{d+2} p_{T_x\mathcal{M}} \right\|_{\mathrm{F}} \leq \left\| \frac{1}{r^2}\Sigma_\mu(x) \right\|_{\mathrm{F}} + \left\| \frac{1}{d+2} p_{T_x\mathcal{M}} \right\|_{\mathrm{F}} \leq 1+1$ to obtain

$$\int_A \left\| \frac{1}{r^2}\Sigma_\mu(x) - \frac{1}{d+2} p_{T_x\mathcal{M}} \right\|_{\mathrm{F}}^p \mathrm{d}\mu(x) \leq 2^p \mu(A).$$

On B, we use Proposition 3.1 to get

$$\int_B \left\| \frac{1}{r^2}\Sigma_\mu(x) - \frac{1}{d+2} p_{T_x\mathcal{M}} \right\|_{\mathrm{F}}^p \mathrm{d}\mu(x) \leq (c_{3.1} r)^p.$$

Combining these two inequalities yields $W_{p,\gamma}^p(\check{\mu}, \check{\mu}_0) \leq \gamma^p (2^p \mu(A) + (c_{3.1} r)^p)$. Using the inequality $(a+b)^{\frac{1}{p}} \leq a^{\frac{1}{p}} + b^{\frac{1}{p}}$, where $a, b \geq 0$, we deduce the result:

$$W_{p,\gamma}(\check{\mu}, \check{\mu}_0) \leq \gamma \left(2\mu(A)^{\frac{1}{p}} + c_{3.1} r \right). \qquad \square$$

3.3 Stability of the estimation

In this subsection we study the stability of the operator $\mu \mapsto \overline{\Sigma}_\mu(\cdot)$ with respect to the W_p metric on measures. The results of this subsection only rely on the following hypotheses about μ:

Hypothesis 5. $\exists c_5 > 0, \forall x \in \mathrm{supp}\,(\mu)$, $\forall t \in [0, \frac{1}{2\rho})$,

$$\mu(\overline{\mathcal{B}}\,(x,t)) \geq c_5 t^d.$$

Hypothesis 6. $\exists c_6 > 0, \forall x \in \mathrm{supp}\,(\mu)$, $\exists \lambda(x) \geq 0$, $\forall s,t \in [0, \lambda(x) \wedge \frac{1}{2\rho})$ s.t. $s \leq t$,

$$\mu(\overline{\mathcal{B}}\,(x,t) \setminus \overline{\mathcal{B}}\,(x,s)) \leq c_6 t^{d-1}(t-s).$$

Hypothesis 7. $\exists c_7 > 0, \forall x \in \mathrm{supp}\,(\mu)$, $\forall s,t \in [0, \frac{1}{2\rho})$ s.t. $s \leq t$,

$$\mu(\overline{\mathcal{B}}\,(x,t) \setminus \overline{\mathcal{B}}\,(x,s)) \leq c_7 t^{d-\frac{1}{2}}(t-s)^{\frac{1}{2}}.$$

Note that, as stated in Propositions 2.17 and 2.18, the inital hypotheses 2 and 3 imply the hypotheses 5, 6 and 7 with $\lambda\,(x)$ being the normal reach of $\mathcal{M}$ at x.

Let μ and ν be two probability measures, $x \in \mathrm{supp}\,(\mu) \cap \mathrm{supp}\,(\nu)$, and consider the Frobenius distance $\left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(x)\right\|_{\mathrm{F}}$ between the normalized local covariance matrices. One shows that this distance is related to the 1-Wasserstein distance between the localized probability measures $\overline{\mu_x}$ and $\overline{\nu_x}$ via the following inequality (see Equation 27 in the proof of Lemma 3.7):

$$\left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(x)\right\|_{\mathrm{F}} \leq \frac{2}{r} W_1(\overline{\mu_x}, \overline{\nu_x}).$$

Without any assumption on the measures, it is not true that $W_1(\overline{\mu_x}, \overline{\nu_x})$ goes to 0 as $W_1(\mu,\nu)$ does. However, if we assume that μ satisfies the hypotheses 5 and 6, that x satisfies $\lambda(x) > 0$ and that r is chosen such that $4\left(\frac{W_1(\mu,\nu)}{c_5 \wedge 1}\right)^{\frac{1}{d+1}} \leq r < \lambda\,(x) \wedge \frac{1}{2\rho}$, then we are able to prove (Lemma C.5) that

$$W_1(\overline{\mu_x}, \overline{\nu_x}) \leq c_{C.5}\left(\frac{W_1(\mu,\nu)}{r^{d-1}}\right)^{\frac{1}{2}}. \tag{25}$$

In Remark C.7, we show that the exponent $d-1$ on r is optimal. As a consequence of this inequality, estimating local covariance matrices is robust in Wasserstein distance:

$$\left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(x)\right\|_{\mathrm{F}} \leq 2c_{C.5}\left(\frac{W_1(\mu,\nu)}{r^{d+1}}\right)^{\frac{1}{2}}. \tag{26}$$

A stability result of this kind already appears in [MSW19, Theorem 4.3], where μ and ν are two probability measures on a bounded set X, and satisfy the following condition: $\forall x \in X, \forall s, r \leq 0$ s.t. $s \leq r$, we have $\frac{\mu(\overline{\mathcal{B}}(x,r))}{\mu(\overline{\mathcal{B}}(x,s))} \leq (\frac{r}{s})^d$. The theorem states that, denoting $D = \mathrm{diam(X)}$, for all $x \in X$,

$$W_1(\overline{\mu_x}, \overline{\nu_x}) \leq (1+2r)\left[\frac{W_1(\mu,\nu)^{\frac{1}{2}}}{1 \wedge (\frac{r}{D})^d} + \left(1 + \frac{W_1(\mu,\nu)^{\frac{1}{2}}}{r}\right)^d - 1\right].$$

When $r \leq D$ and $W_1(\mu,\nu)$ goes to zero, we obtain that $W_1(\overline{\mu_x}, \overline{\nu_x})$ is of order

$$W_1(\overline{\mu_x}, \overline{\nu_x}) \leq (1+2r)D^d\left(\frac{W_1(\mu,\nu)}{r^{2d}}\right)^{\frac{1}{2}}.$$

The exponent on r is greater here than in Equation 25.

Another result in [MMM18, Theorem 3] bounds the distance $\left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(x)\right\|_{\mathrm{F}}$ with the ∞-Wasserstein distance $W_\infty(\mu, \nu)$. Namely, if μ and ν are fully supported probability measures with densities upper bounded by $l > 0$ and supports included in $X \subset \mathbb{R}^d$, denoting $D = \mathrm{diam(X)}$, we have

$$\left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(x)\right\|_{\mathrm{F}} \leq lAW_\infty(\mu, \nu),$$

where $A = \frac{d}{d+2}\frac{(r+D)^{d+1}}{Dr^d} + \frac{(2r+D)(r+D)^d}{r^d} + \frac{2d}{d+2}\frac{(r+D)^{d+2}}{Dr^d}$.

Remark 3.5. Let us show that in general, for $x \in \mathrm{supp}\,(\mu) \cap \mathrm{supp}\,(\nu)$, it is not true that $\left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(x)\right\|_{\mathrm{F}}$ goes to zero as $W_1(\mu, \nu)$ goes to zero. Similarly, $W_{p,\gamma}(\check{\mu}, \check{\nu})$ does not have to go to zero. For example, one can consider $\epsilon > 0$, and the measures on $\mathbb{R}$

$$\mu = \tfrac{1}{2}(\delta_0 + \delta_1) \qquad \text{and} \qquad \nu = \tfrac{1}{2}(\delta_0 + \delta_{1+\epsilon}).$$

Choose the scale parameter $r = 1$. We have $\Sigma_\mu(0) = \Sigma_\mu(1) = \frac{1}{2}1^{\otimes 2}$ and $\Sigma_\nu(0) = \Sigma_\nu(1+\epsilon) = 0$. The measures $\check{\mu}$ and $\check{\nu}$ on $\mathbb{R} \times \mathcal{M}(\mathbb{R})$ can be written

$$\check{\mu} = \tfrac{1}{2}\left(\delta_{\left(0, \frac{1}{2}1^{\otimes 2}\right)} + \delta_{\left(1, \frac{1}{2}1^{\otimes 2}\right)}\right) \qquad \text{and} \qquad \check{\nu} = \tfrac{1}{2}\left(\delta_{(0,0)} + \delta_{(1+\epsilon,0)}\right).$$

A computation shows that

$$\begin{aligned} W_{p,\gamma}^p(\check{\mu}, \check{\nu}) &= \frac{1}{2}\left\|\left(0, \frac{1}{2}1^{\otimes 2}\right) - \left(0, 0\right)\right\|_\gamma^p + \frac{1}{2}\left\|\left(1, \frac{1}{2}1^{\otimes 2}\right) - \left(1+\epsilon, 0\right)\right\|_\gamma^p \\ &= \frac{1}{2}\left(\left(\frac{\gamma}{2}\right)^p + \left(\epsilon^2 + \gamma^2\frac{1}{4}\right)^{\frac{p}{2}}\right) \geq \left(\frac{\gamma}{2}\right)^p. \end{aligned}$$

Hence $W_{p,\gamma}(\check{\mu}, \check{\nu}) \geq \frac{\gamma}{2} > 0$. Besides, we have $W_1(\mu, \nu) = \frac{1}{2}\epsilon$. Hence $W_{p,\gamma}(\check{\mu}, \check{\nu})$ does not go to zero as $W_1(\mu, \nu)$ does. However, under regularity assumptions on μ, the following proposition states that it is the case.

Proposition 3.6. *Let μ and ν be two probability measures on E. Suppose that μ statisfies the hypotheses 5, 6 and 7. Define $w = W_p(\mu, \nu)$. Suppose that $r \leq \frac{1}{2\rho} \wedge 1$ and $w \leq (c_5 \wedge 1)(\frac{r}{4})^{d+1}$. Then*

$$W_{p,\gamma}(\check{\mu}, \check{\nu}) \leq 2w + \gamma c_{3.6}\left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}} + \gamma c_{3.6}'\mu(\lambda^r)^{\frac{1}{p}}\left(\frac{w}{r^{d+1}}\right)^{\frac{1}{4}}$$

with $c_{3.6} = 4(1 + c_{3.8})$ and $c_{3.6}' = 4c_{C.6}$.

Proof. According to Lemma 3.7 stated below, we have

$$W_{p,\gamma}(\check{\mu}, \check{\nu}) \leq 2^{\frac{p-1}{p}}\left(1 + \frac{2\gamma}{r}\right)w + 2^{\frac{p-1}{p}}\frac{2\gamma}{r}\left(\int W_1^p(\overline{\mu_x}, \overline{\nu_y})\mathrm{d}\pi(x, y)\right)^{\frac{1}{p}}.$$

Let $\alpha = \left(\frac{w}{r^{d-1}}\right)^{\frac{1}{2}}$. Lemma 3.8, also stated below, gives

$$\left(\int W_1(\overline{\mu_x}, \overline{\nu_y})\mathrm{d}\pi(x, y)\right)^{\frac{1}{p}} \leq 2^{\frac{p-1}{p}}\left(c_{C.6}r^{\frac{1}{2}}\mu(\lambda^r)^{\frac{1}{p}}\alpha^{\frac{1}{2}} + c_{3.8}\alpha\right)$$

Combining these inequalities yields

$$\begin{aligned} W_{p,\gamma}(\check{\mu}, \check{\nu}) &\leq 2^{\frac{p-1}{p}}w + 2^{\frac{p-1}{p}}\frac{2\gamma}{r}\left(w + 2^{\frac{p-1}{p}}c_{3.8}\alpha\right) + \left(2^{\frac{p-1}{p}}\right)^2\frac{2\gamma}{r}c_{C.6}r^{\frac{1}{2}}\mu(\lambda^r)^{\frac{1}{p}}\alpha^{\frac{1}{2}} \\ &\leq 2w + 2\cdot 2\gamma\left(\frac{w}{r} + 2c_{3.8}\frac{\alpha}{r}\right) + 2^2\cdot 2\gamma c_{C.6}\mu(\lambda^r)^{\frac{1}{p}}\left(\frac{\alpha}{r}\right)^{\frac{1}{2}}, \end{aligned}$$

where we used $2^{\frac{p-1}{p}} \le 2$. Since $r \le 1$, we have $w \le 1$ and $w = \left(\frac{w}{r^{d-1}}\right)^{\frac{1}{2}} r^{\frac{d-1}{2}} w^{\frac{1}{2}} \le \left(\frac{w}{r^{d-1}}\right)^{\frac{1}{2}} = \alpha$. We get

$$W_{p,\gamma}(\check{\mu},\check{\nu}) \le 2^{\frac{p-1}{p}} w + 2^{\frac{p-1}{p}} 2\gamma \left(1 + 2^{\frac{p-1}{p}} c_{3.8}\right) \frac{\alpha}{r} + \left(2^{\frac{p-1}{p}}\right)^2 2\gamma c_{C.6} \mu(\lambda^r)^{\frac{1}{p}} \left(\frac{\alpha}{r}\right)^{\frac{1}{2}}.$$

By replacing $\frac{\alpha}{r}$ with $\left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}}$, we obtain the result. □

Let us interpret the inequality

$$W_{p,\gamma}(\check{\mu},\check{\nu}) \le 2w + \gamma c_{3.6} \left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}} + \gamma c'_{3.6} \mu(\lambda^r)^{\frac{1}{p}} \left(\frac{w}{r^{d+1}}\right)^{\frac{1}{4}}$$

The first term $2w$ is to be seen as the initial error between the measures μ and ν. The second term $\gamma c_{3.6}(\frac{w}{r^{d+1}})^{\frac{1}{2}}$ corresponds to the local errors $W_1(\overline{\mu_x}, \overline{\nu_y})$ when comparing the normalized covariance matrices. The third term $\gamma c'_{3.6} \mu(\lambda^r)^{\frac{1}{p}} \left(\frac{w}{r^{d+1}}\right)^{\frac{1}{4}}$ stands for the error on points x such that $\lambda(x) \le r$, where the stability is weaker.

As a consequence of this proposition, the application $\mu \mapsto \check{\mu}$, seen as an application between spaces of measures endowed with the Wassertein metric, is continuous on the set of measures μ which satisfy 5, 6 and 7 with $\frac{1}{2\rho} \ge r$.

We now state the lemmas used in the proof of Proposition 3.6.

Lemma 3.7. *Let π be an optimal transport plan for $W_p(\mu,\nu)$. Then*

$$W_{p,\gamma}(\check{\mu},\check{\nu}) \le 2^{\frac{p-1}{p}} \left(1 + \frac{2\gamma}{r}\right) W_p(\mu,\nu) + 2^{\frac{p-1}{p}} \frac{2\gamma}{r} \left(\int W_1^p(\overline{\mu_x}, \overline{\nu_y}) \mathrm{d}\pi(x,y)\right)^{\frac{1}{p}}.$$

Proof. We first prove the following fact: for every $x \in \operatorname{supp}(\mu)$ and $y \in \operatorname{supp}(\nu)$,

$$\|\Sigma_\mu(x) - \Sigma_\nu(y)\|_{\mathrm{F}} \le 2r\left(\|x-y\| + W_1(\overline{\mu_x}, \overline{\nu_y})\right). \tag{27}$$

Let ρ be any transport plan between $\overline{\mu_x}$ and $\overline{\nu_y}$. We have

$$\begin{aligned} \Sigma_\mu(x) - \Sigma_\nu(y) &= \int (x-y)^{\otimes 2} \mathrm{d}\overline{\mu_x}(x') - \int (y-y')^{\otimes 2} \mathrm{d}\overline{\mu_y}(y') \\ &= \int \left((x-x')^{\otimes 2} - (y-y')^{\otimes 2}\right) \mathrm{d}\rho(x',y'). \end{aligned} \tag{28}$$

For any $x' \in \overline{\mathcal{B}}(x,r)$ and $y' \in \overline{\mathcal{B}}(y,r)$, we can use Lemma C.1 to get

$$\left\|(x-x')^{\otimes 2} - (y-y')^{\otimes 2}\right\|_{\mathrm{F}} \le (r+r)(\|x-y\| + \|x'-y'\|).$$

Therefore, Equation 28 yields

$$\begin{aligned} \|\Sigma_\mu(x) - \Sigma_\nu(y)\|_{\mathrm{F}} &\le \int 2r(\|x-y\| + \|x'-y'\|) \mathrm{d}\rho(x',y') \\ &\le 2r\left(\|x-y\| + W_1(\overline{\mu_x}, \overline{\nu_y})\right). \end{aligned}$$

Now, a transport plan π for $W_p(\mu,\nu)$ begin given, we build a transport plan $\check{\pi}$ for $(\check{\mu},\check{\nu})$ as follows: for every $\phi\colon (E \times \mathcal{M}(E))^2 \to \mathbb{R}$ with compact support, let $\check{\pi}$ satisfies

$$\int \phi(x,A,y,B) \mathrm{d}\check{\pi}(x,A,y,B) = \int \phi\left(x, \overline{\Sigma}_\mu(x), y, \overline{\Sigma}_\nu(y)\right) \mathrm{d}\pi(x,y).$$

We have the majoration

$$
\begin{aligned}
W_{p,\gamma}^p(\check{\mu},\check{\nu}) &\leq \int \|(x,A)-(y,B)\|_\gamma^p \,\mathrm{d}\check{\pi}(x,A,y,B) \\
&= \int \left(\|x-y\|^2 + \gamma^2 \left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(y)\right\|_{\mathrm{F}}^2\right)^{\frac{p}{2}} \mathrm{d}\pi(x,y) \\
&\leq \int \left(\|x-y\| + \gamma \left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(y)\right\|_{\mathrm{F}}\right)^p \mathrm{d}\pi(x,y) \qquad (29)
\end{aligned}
$$

Besides, Equation 27 gives

$$
\left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(y)\right\|_{\mathrm{F}} \leq \frac{1}{r^2}\left\|\Sigma_\mu(x) - \Sigma_\nu(y)\right\|_{\mathrm{F}} \leq \frac{2}{r}\left(\|x-y\| + W_1(\overline{\mu_x},\overline{\nu_y})\right).
$$

We can use the inequality $(a+b)^p \leq 2^{p-1}(a^p+b^p)$, where $a,b\geq 0$, to deduce

$$
\begin{aligned}
\left(\|x-y\| + \gamma \left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(y)\right\|_{\mathrm{F}}\right)^p &\leq \left(\|x-y\| + \gamma\frac{2}{r}\left(\|x-y\| + W_1(\overline{\mu_x},\overline{\nu_y})\right)\right)^p \\
&\leq 2^{p-1}\left(\left(1+\frac{2\gamma}{r}\right)\|x-y\|\right)^p + 2^{p-1}\left(\frac{2\gamma}{r}W_1(\overline{\mu_x},\overline{\nu_y}\right)^p
\end{aligned}
$$

By inserting this inequality in Equation 29 we obtain

$$
\begin{aligned}
W_{p,\gamma}^p(\check{\mu},\check{\nu}) &\leq 2^{p-1}\int \left(\left(1+\frac{2\gamma}{r}\right)\|x-y\|\right)^p + \left(\frac{2\gamma}{r}W_1(\overline{\mu_x},\overline{\nu_y})\right)^p \mathrm{d}\pi(x,y) \\
&= 2^{p-1}\left(1+\frac{2\gamma}{r}\right)^p W_p^p(\mu,\nu) + 2^{p-1}\left(\frac{2\gamma}{r}\right)^p \int W_1^p(\overline{\mu_x},\overline{\nu_y})\mathrm{d}\pi(x,y),
\end{aligned}
$$

which yields the result. □

Lemma 3.8. *Let $w = W_p(\mu,\nu)$ and define $\alpha = (\frac{w}{r^{d-1}})^{\frac{1}{2}}$. Suppose that $r \leq \frac{1}{2\rho}$ and $w \leq (c_5 \wedge 1)(\frac{r}{4})^{d+1}$. Let π be an optimal transport plan for $W_p(\mu,\nu)$. Then*

$$
\begin{aligned}
&\left(\int W_1^p(\overline{\mu_x},\overline{\nu_y})\mathrm{d}\pi(x,y)\right)^{\frac{1}{p}} \\
&\qquad \leq 2^{\frac{p-1}{p}}\left(c_{C.6} r^{\frac{1}{2}} \mu(\lambda^r)^{\frac{1}{p}}\alpha^{\frac{1}{2}} + \left(2r^d + c_{C.4} r^{\frac{d+1}{2}} + c_{C.5}\right)\alpha + (1+c_{C.3})w\right).
\end{aligned}
$$

If we suppose that $r \leq 1$, then

$$
\left(\int W_1^p(\overline{\mu_x},\overline{\nu_y})\mathrm{d}\pi(x,y)\right)^{\frac{1}{p}} \leq 2^{\frac{p-1}{p}}\left(c_{C.6} r^{\frac{1}{2}}\mu(\lambda^r)^{\frac{1}{p}}\alpha^{\frac{1}{2}} + c_{3.8}\alpha\right)
$$

with $c_{3.8} = 3 + c_{C.3} + c_{C.4} + c_{C.5}$.

Proof. We denote $w = W_p(\mu,\nu)$ and $\alpha = (\frac{w}{r^{d-1}})^{\frac{1}{2}}$. Let us cut the integral as follows:

$$
\int W_1^p(\overline{\mu_x},\overline{\nu_y})d\pi(x,y) = \int_A + \int_B + \int_C W_1^p(\overline{\mu_x},\overline{\nu_y})\mathrm{d}\pi(x,y)
$$

where $A = \{(x,y), \|x-y\| \geq \alpha\}$, $B = \{(x,y), \|x-y\| < \alpha \text{ and } \lambda(x) > r\}$ and $C = \{(x,y), \|x-y\| < \alpha \text{ and } \lambda(x) \leq r\}$.

Term A: We use the following loose majoration:

$$
\begin{aligned}
W_1(\overline{\mu_x},\overline{\nu_y}) &\leq W_1(\overline{\mu_x},\delta_x) + W_1(\delta_x,\delta_y) + W_1(\delta_y,\overline{\nu_y}) \\
&\leq r + \|x-y\| + r
\end{aligned}
$$

to obtain $W_1^p(\overline{\mu_x}, \overline{\nu_y}) \leq 2^{p-1}\big((2r)^p + \|x-y\|^p\big)$ and

$$\begin{aligned}
\int_A W_1^p(\overline{\mu_x}, \overline{\nu_y}) \mathrm{d}\pi(x,y) &\leq \int_A 2^{p-1}\big((2r)^p + \|x-y\|^p\big) \mathrm{d}\pi(x,y) \\
&\leq 2^{p-1}(2r)^p \pi(A) + \int 2^{p-1} \|x-y\|^p \mathrm{d}\pi(x,y) \\
&= 2^{p-1}(2r)^p \pi(A) + 2^{p-1} w^p.
\end{aligned}$$

But $\pi(A) = \pi(\{(x,y), \|x-y\| > \alpha) = \pi\left(\{(x,y), \|x-y\|^p > \alpha^p)\}\right) \leq \left(\frac{w}{a}\right)^p$ by Markov inequality. Therefore,

$$\begin{aligned}
\int_A W_1^p(\overline{\mu_x}, \overline{\nu_y}) \mathrm{d}\pi(x,y) &\leq 2^{p-1}(2r)^p \left(\frac{w}{\alpha}\right)^p + 2^{p-1} w^p \\
&= 2^{p-1}(2r^d\alpha)^p + 2^{p-1} w^p,
\end{aligned}$$

where we used $r\frac{w}{\alpha} = r^d\alpha$ on the last line.

Term B: On the event B, we write

$$W_1(\overline{\mu_x}, \overline{\nu_y}) \leq W_1(\overline{\mu_x}, \overline{\mu_y}) + W_1(\overline{\mu_y}, \overline{\nu_y}).$$

Since $\lambda(x) > r$, Lemma C.3 and Lemma C.5 give $W_1(\overline{\mu_x}, \overline{\mu_y}) \leq c_{C.3}\|x-y\|$ and $W_1(\overline{\mu_y}, \overline{\nu_y}) \leq c_{C.5}\alpha$. We deduce that

$$\begin{aligned}
\int_B W_1^p(\overline{\mu_x}, \overline{\nu_y}) \mathrm{d}\pi(x,y) &\leq 2^{p-1} \int_B (c_{C.3}\|x-y\|)^p + (c_{C.5}\alpha)^p \mathrm{d}\pi(x,y) \\
&\leq 2^{p-1}(c_{C.3}w)^p + 2^{p-1}(c_{C.5}\alpha)^p.
\end{aligned}$$

Term C: We proceed as for Term B, but using Lemmas C.4 and C.6 instead of Lemmas C.3 and C.5. This yields

$$\begin{aligned}
W_1(\overline{\mu_x}, \overline{\nu_y}) &\leq W_1(\overline{\mu_x}, \overline{\mu_y}) + W_1(\overline{\mu_y}, \overline{\nu_y}) \\
&\leq c_{C.4} r^{\frac{1}{2}} \|x-y\|^{\frac{1}{2}} + c_{C.6} r^{\frac{1}{2}} \alpha^{\frac{1}{2}},
\end{aligned}$$

and we deduce that

$$\int_C W_1^p(\overline{\mu_x}, \overline{\nu_y}) \mathrm{d}\pi(x,y) \leq \int_C 2^{p-1}\left(c_{C.4} r^{\frac{1}{2}} \|x-y\|^{\frac{1}{2}}\right)^p \mathrm{d}\pi(x,y) + 2^{p-1}\pi(C)\left(c_{C.6} r^{\frac{1}{2}} \alpha^{\frac{1}{2}}\right)^p. \tag{30}$$

On the one hand, we have $\int_C \|x-y\|^{\frac{p}{2}} \mathrm{d}\pi(x,y) \leq \int_{E\times E} \|x-y\|^{\frac{p}{2}} \mathrm{d}\pi(x,y)$, and by Jensen's inequality,

$$\int_{E\times E} \|x-y\|^{\frac{p}{2}} \mathrm{d}\pi(x,y) \leq (w^p)^{\frac{1}{2}}.$$

On the other hand, by definition of C, we have $\pi(C) \leq \mu(\lambda^r)$. Combined with Equation 30, we obtain

$$\int_C W_1(\overline{\mu_x}, \overline{\nu_y}) \mathrm{d}\pi(x,y) \leq 2^{p-1}\left(c_{C.4} r^{\frac{1}{2}} w^{\frac{1}{2}}\right)^p + 2^{p-1}\mu(\lambda^r)\left(c_{C.6} r^{\frac{1}{2}} \alpha^{\frac{1}{2}}\right)^p.$$

To conclude the proof, we write

$$\begin{aligned}
\int W_1(\overline{\mu_x}, \overline{\nu_y}) \mathrm{d}\pi(x,y) &= \int_A + \int_B + \int_C W_1(\overline{\mu_x}, \overline{\nu_y}) \mathrm{d}\pi(x,y) \\
&\leq 2^{p-1}(2r^d\alpha)^p + 2^{p-1}w^p + 2^{p-1}(c_{C.3}w)^p + 2^{p-1}(c_{C.5}\alpha)^p \\
&\quad + 2^{p-1}\left(c_{C.4} r^{\frac{1}{2}} w^{\frac{1}{2}}\right)^p + 2^{p-1}\mu(\lambda^r)\left(c_{C.6} r^{\frac{1}{2}} \alpha^{\frac{1}{2}}\right)^p.
\end{aligned}$$

We use the inequality $(a+b)^{\frac{1}{p}} \leq a^{\frac{1}{p}} + b^{\frac{1}{p}}$, where $a, b \geq 0$, to get

$$
\begin{aligned}
&\left(\int W_1(\overline{\mu_x}, \overline{\nu_y}) \mathrm{d}\pi(x,y)\right)^{\frac{1}{p}} \\
&\leq 2^{\frac{p-1}{p}} \left(2r^d\alpha + w + c_{C.3}w + c_{C.5}\alpha + c_{C.4}r^{\frac{1}{2}}w^{\frac{1}{2}} + \mu(\lambda^r)^{\frac{1}{p}} c_{C.6} r^{\frac{1}{2}}\alpha^{\frac{1}{2}}\right) \\
&\leq 2^{\frac{p-1}{p}} \left(c_{C.6} r^{\frac{1}{2}} \mu(\lambda^r)^{\frac{1}{p}} \alpha^{\frac{1}{2}} + \left(2r^d + c_{C.4} r^{\frac{d+1}{2}} + c_{C.5}\right)\alpha + (1 + c_{C.3})w\right)
\end{aligned}
$$

where we used $c_{C.4}r^{\frac{1}{2}}w^{\frac{1}{2}} = c_{C.4}r^{\frac{d+1}{2}}\alpha$ on the the last line. This proves the first result.

If we suppose $r \leq 1$, we can use the inequalities $r^d \leq r^{\frac{d+1}{2}} \leq 1$ and $w = \alpha r^{\frac{d-1}{2}} w^{\frac{1}{2}} \leq \alpha$ to obtain the simplified expression

$$
\left(\int W_1(\overline{\mu_x}, \overline{\nu_y}) \mathrm{d}\pi(x,y)\right)^{\frac{1}{p}} \leq 2^{\frac{p-1}{p}} \left(c_{C.6} r^{\frac{1}{2}} \mu(\lambda^r)^{\frac{1}{p}} \alpha^{\frac{1}{2}} + (3 + c_{C.3} + c_{C.4} + c_{C.5})\,\alpha\right) \qquad \square
$$

Remark 3.9. On Term C, we could have used the inequality $W_1(\overline{\mu_x}, \overline{\nu_y}) \leq r + \|x-y\| + r$ to obtain

$$
\begin{aligned}
\int_C W_1^p(\overline{\mu_x}, \overline{\nu_y}) \mathrm{d}\pi(x,y) &\leq 2^{p-1} \int_C (2r)^p + \|x-y\|^p \,\mathrm{d}\pi(x,y) \\
&\leq 2^{p-1}(2r)^p \pi(C) + 2^{p-1} w^p.
\end{aligned}
$$

Following the rest of the proof, and under the assumption $r \leq 1$, we eventually obtain

$$
\left(\int W_1(\overline{\mu_x}, \overline{\nu_y}) \mathrm{d}\pi(x,y)\right)^{\frac{1}{p}} \leq 2^{\frac{p-1}{p}} \left(2r\mu(\lambda^r)^{\frac{1}{p}} + c'_{3.8}\alpha\right)
$$

with $c'_{3.8} = 4 + c_{C.3} + c_{C.5}$.

Note that in the term $r\mu(\lambda^r)^{\frac{1}{p}}$, the exponent over r is better than in Lemma 3.8, which is $r^{\frac{1}{2}}\mu(\lambda^r)^{\frac{1}{p}}\alpha^{\frac{1}{2}}$. However, we prefer to keep the term $\alpha^{\frac{1}{2}}$, for it goes to zero as w does.

3.4 An approximation theorem

Let us recall the definitions of Subsection 3.1: the exact lifted measure is $\check{\mu}_0 = (u_*\mu_0)(x_0) \otimes \delta_{\frac{1}{d+2}p_{T_x\mathcal{M}}}$, and the lifted measure associated to ν is $\check{\nu} = \nu(x) \otimes \delta_{\overline{\Sigma}_\nu(x)}$. We are now able to state that $\check{\nu}$ is close to $\check{\mu}_0$, that is, $\check{\nu}$ is a consistent estimator of $\check{\mu}_0$, in Wasserstein distance.

Theorem 3.10. *Assume that $\mathcal{M}_0$ and μ_0 satisfy the hypotheses 1, 2, 3. Let ν be any probability measure. Denote $w = W_p(\mu, \nu)$. Suppose that $r \leq \frac{1}{2\rho} \wedge 1$ and $w \leq (c_5 \wedge 1)(\frac{r}{4})^{d+1}$. Then*

$$
W_{p,\gamma}(\check{\nu}, \check{\mu}_0) \leq \gamma c_{3.10}\mu(\lambda^r)^{\frac{1}{p}} + \gamma c_{3.1} r + \gamma c_{3.6}\left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}} + 2w
$$

where $c_{3.10} = 2 + \frac{1}{2}c'_{3.6}$.

Proof. It is a direct consequence of Propositions 3.4 and 3.6. $\square$

In order to simplify the results of the following section, we shall use a weaker result. Using Hypothesis 4, we get rid of the term $\mu(\lambda^r)$.

Corollary 3.11. *Let $r > 0$. Assume that $\mathcal{M}_0$ and μ_0 satisfy the hypotheses 1, 2, 3 and Hypothesis 4 with $r_4 \geq r$. Let ν be any probability measure. Denote $w = W_p(\mu, \nu)$. Suppose that $r \leq \frac{1}{2\rho} \wedge 1$ and $w \leq (c_5 \wedge 1)(\frac{r}{4})^{d+2}$. Then*

$$
W_{p,\gamma}(\check{\nu}, \check{\mu}_0) \leq \big(1 + \gamma c_{3.11}\big) r^{\frac{1}{p}}
$$

with $c_{3.11} = c_{3.10}(c_4)^{\frac{1}{p}} + c_{3.6} + c_{3.1}$.

Proof. According to Theorem 3.10, we have

$$W_{p,\gamma}(\check{\nu}, \check{\mu}_0) \leq \gamma c_{3.10} \mu(\lambda^r)^{\frac{1}{p}} + \gamma c_{3.1} r + \gamma c_{3.6} \left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}} + 2w.$$

Note that the assumption $w \leq (c_5 \wedge 1)(\frac{r}{4})^{d+2}$ yields $\left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}} \leq \frac{r}{4} \leq r$. Besides, $r \leq 1$ yields $w \leq \left(\frac{r}{4}\right)^{d+2} \leq \frac{r}{4} \leq \frac{r}{2}$. Finally, Hypothesis 4 gives $\mu(\lambda^r) \leq c_4 r$, and we deduce the result thanks to the rough majoration $r \leq r^{\frac{1}{p}}$:

$$\begin{aligned} W_{p,\gamma}(\check{\nu}, \check{\mu}_0) &\leq \gamma c_{3.10} (c_4 r)^{\frac{1}{p}} + \gamma c_{3.1} r + \gamma c_{3.6} r + r \\ &\leq \left(\gamma c_{3.10} (c_4)^{\frac{1}{p}} + \gamma c_{3.1} + \gamma c_{3.6} + 1\right) r^{\frac{1}{p}}. \end{aligned}$$

□

4 Topological inference with the lifted measure

Based on the results of the last section, we show how the lifted measure $\check{\nu}$ can be used to infer the homotopy type of $\check{\mathcal{M}}$, or to estimate the persistent homology of $\check{\mu}_0$.

4.1 Overview of the method

Let us recall the results obtained so far. Assume that the immersion $u\colon \mathcal{M}_0 \to \mathcal{M}$ and the measure μ_0 satisfy the Hypotheses 1, 2 and 3. Our goal is to estimate the exact lifted measure $\check{\mu}_0$ on $E \times \mathcal{M}(E)$, since its support is the submanifold $\check{\mathcal{M}}$, which is diffeomorphic to $\mathcal{M}_0$.

To do so, we suppose that we are observing a measure ν on E. No assumptions are made on ν. Our results only depends on the Wasserstein distance

$$w = W_p(\mu, \nu),$$

where $\mu = u_* \mu_0$. Recall that the measure $\check{\mu}_0$ is defined as (Equation 20):

$$\check{\mu}_0 = (u_* \mu_0)(x_0) \otimes \delta_{\frac{1}{d+2} p_{T_x \mathcal{M}}}.$$

To approximate $\check{\mu}_0$, we pick a parameter $r > 0$ and consider the lifted measure $\check{\nu}$ built on ν (Definition 3.2):

$$\check{\nu} = \nu(x) \otimes \delta_{\overline{\Sigma}_\nu(x)}.$$

Choose $\gamma > 0$. Endow the space $E \times \mathcal{M}(E)$ with the norm $\|\cdot\|_\gamma$ (Equation 21), and consider the Wasserstein distance $W_{p,\gamma}(\cdot, \cdot)$ between measures on $E \times \mathcal{M}(E)$ (Equation 22). We quantify the quality of the approximation by the Wasserstein distance

$$W_{p,\gamma}(\check{\mu}_0, \check{\nu}).$$

According to Theorem 3.10, we have

$$W_{p,\gamma}(\check{\nu}, \check{\mu}_0) \leq \gamma c_{3.10} \mu(\lambda^r)^{\frac{1}{p}} + \gamma c_{3.1} r + \gamma c_{3.6} \left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}} + 2w$$

as long as the parameter r satisfies

$$4\left(\frac{w}{c_5 \wedge 1}\right)^{\frac{1}{d+1}} \;\leq\; r \;\leq\; \frac{1}{2\rho} \wedge 1.$$

Under Hypothesis 4, Corollary 3.11 gives a weaker form of this result. We have

$$W_{p,\gamma}(\check{\nu}, \check{\mu}_0) \leq \big(1 + \gamma c_{3.11}\big) r^{\frac{1}{p}}$$

as long as the parameter r satisfies

$$4\left(\frac{w}{c_5 \wedge 1}\right)^{\frac{1}{d+2}} \leq r \leq \frac{1}{2\rho} \wedge r_4 \wedge 1.$$

In the following subsections, we show how these results lead to consistent estimations of $\mathcal{M}_0$ and its homology. Namely, we can estimate the homotopy type of $\check{\mathcal{M}}$, and hence of $\mathcal{M}_0$, by considering the sublevel sets of the DTM $d_{\check{\nu},m,\gamma}$ (Corollary 4.3). The notation $d_{\check{\nu},m,\gamma}$ corresponds to the DTM, defined in Subsection 1.4, seen in the ambient space $\left(E \times \mathcal{M}(E), \|\cdot\|_\gamma\right)$. Besides, we can estimate the persistent homology of the DTM-filtration $W_\gamma[\check{\mu}_0]$ with the filtration $W_\gamma[\check{\nu}]$ (Corollary 4.5). Here, $W_\gamma[\cdot]$ corresponds to the DTM-filtration in the ambient space $\left(E \times \mathcal{M}(E), \|\cdot\|_\gamma\right)$.

Example 4.1. Let $\mathcal{M}$ be the lemniscate of Bernoulli of diameter 2. It is the immersion of a circle $\mathcal{M}_0$. We observe a 100-sample X of $\mathcal{M}$ (Figure 21). Experimentally, we computed the Hausdorff distance $d_{\mathrm{H}}(\mathcal{M}, X) \approx 0{,}026$. Let μ be the Hausdorff measure on $\mathcal{M}$ and ν the empirical measure on X. We choose the parameter $p = 2$. Their Wasserstein distance is approximately $W_2(\mu, \nu) \approx 0{,}015$.

Figure 21: Left: The lemniscate $\mathcal{M}$. Right: The set X, a 100-sample of $\mathcal{M}$.

For each point x of X, we compute the matrix $\overline{\Sigma}_\nu(x)$ with parameter $r = 0{,}5$ and $0{,}1$. This matrix is used as an estimator of the tangent space $T_x\mathcal{M}$. In order the observe the quality of this estimation, we represent on Figure 22 (first row) the principal axes of $\overline{\Sigma}_\nu(x)$ for some x. On the second row are represented the distances $\left\|\overline{\Sigma}_\nu(x) - \frac{1}{d+2}p_{T_x\mathcal{M}}\right\|_{\mathrm{F}}$. One sees that $r = 0{,}1$ yields a better approximation. However, the estimation is still biaised next to the self-intersection points of $\mathcal{M}$.

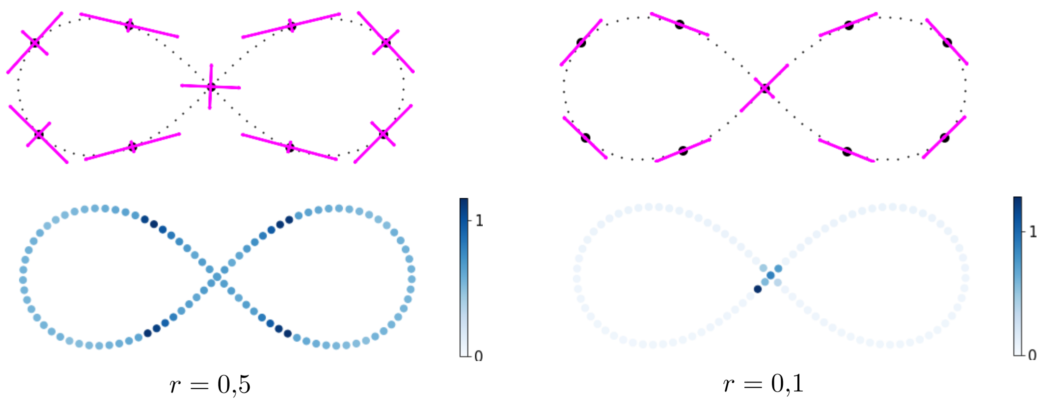

Figure 22: First row: The eigenvectors of $\overline{\Sigma}_\nu(x)$ for some $x \in X$, weighted with their corresponding eigeinvalue. Second row: color representation of the distances $\left\|\overline{\Sigma}_\nu(x) - \frac{1}{d+2}p_{T_x\mathcal{M}}\right\|_{\mathrm{F}}$.

Now we choose the parameter $\gamma = 2$. For $r = 0{,}5$ and $0{,}1$, we consider the lifted measures built on ν, repectively denoted $\check{\nu}^{0,5}$ and $\check{\nu}^{0,1}$. They are measure on the lift space $\mathbb{R}^2 \times \mathcal{M}(\mathbb{R}^2)$, which is endowed with the norm $\|\cdot\|_\gamma$. We computed the Wasserstein distances:

$$W_{2,\gamma}\left(\check{\mu}_0, \check{\nu}^{0,5}\right) \approx 0{,}674 \qquad \text{and} \qquad W_{2,\gamma}\left(\check{\mu}_0, \check{\nu}^{0,1}\right) \approx 0{,}200.$$

In comparison, even with a small parameter r, the Hausdorff distance between their support is still large:

$$d_{\mathrm{H}}\left(\check{\mathcal{M}}, \operatorname{supp}\left(\check{\nu}^{0,5}\right)\right) \approx 1{,}142 \qquad \text{and} \qquad d_{\mathrm{H}}\left(\check{\mathcal{M}}, \operatorname{supp}\left(\check{\nu}^{0,1}\right)\right) \approx 1{,}273.$$

These sets are represented in Figure 23. Observe that, at the center of the graphs, the measures $\check{\nu}^{0,5}$ and $\check{\nu}^{0,1}$ deviate from the set $\check{\mathcal{M}}$.

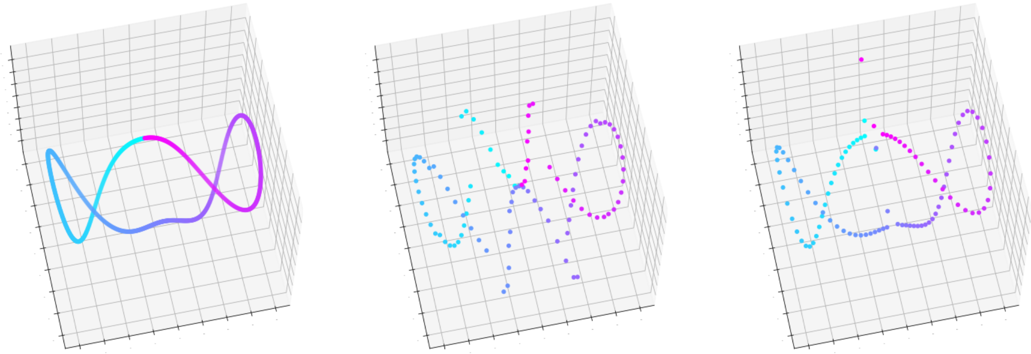

Figure 23: Left: The lifted lemniscate $\check{\mathcal{M}}$, projected in a 3-dimensional subspace via PCA. Center: The set $\operatorname{supp}\left(\check{\nu}^{0,5}\right)$ projected in the same 3-dimensional subspace. Right: Same for $\operatorname{supp}\left(\check{\nu}^{0,1}\right)$.

Example 4.2. Let $u\colon \mathcal{M}_0 \to \mathcal{M}$ be the figure-8 immersion of the torus in $\mathbb{R}^3$, represented in Figure 24. It can be parametrized by rotating a lemniscate around an axis, while forming a full twist. The self-intersection points of this immersion corresponds to the inner circle formed by the center of the lemniscate. These are the points x of $\mathcal{M}$ such that their normal reach $\lambda(x)$ is zero.

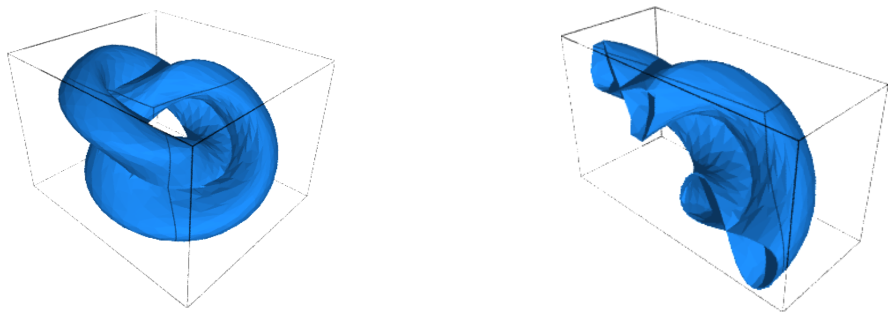

Figure 24: Left: The immersion $\mathcal{M}$ of the torus. Right: A section of $\mathcal{M}$. One sees the inner lemniscate.

Let $\check{\mathcal{M}}$ be the lift of $\mathcal{M}_0$. It is a submanifold of $\mathbb{R}^3 \times \mathcal{M}(\mathbb{R}^3)$. One cannot embed $\check{\mathcal{M}}$ in $\mathbb{R}^3$ by performing a PCA. However, we can try to visualize $\check{\mathcal{M}}$ by considering a small section of it. Figure 25 represents a subset of $\check{\mathcal{M}}$, projected in a 3-dimensional subspace via PCA. One sees that it does not self-intersect.

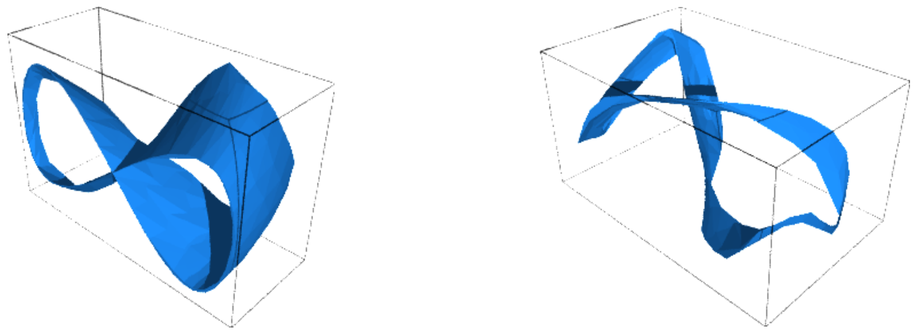

Figure 25: Left: A section of $\mathcal{M}$. Right: The corresponding section of $\check{\mathcal{M}}$, projected in a 3-dimensional subspace via PCA. Observe that it does not self-intersect.

In order to fit in the context of our study, let μ be the Hausdorff measure on $\mathcal{M}$. We observe a 9000-sample X of $\mathcal{M}$, and consider its empirical measure ν. The set X is depicted in Figure 26. Choose the parameter $p = 1$. We compute the Wasserstein distance $W_1(\mu, \nu) \approx 0{,}070$ and the Hausdorff distance $d_{\mathrm{H}}(\mathcal{M}, X) = 0{,}083$.

Let $r = 0{,}09$. In order to observe the estimation of tangent spaces by local covariance matrices $\overline{\Sigma}_\nu(x)$ with parameter r, we represent on Figure 26 the points x such that the distance $\left\|\overline{\Sigma}_\nu(x) - \frac{1}{d+2} p_{T_x\mathcal{M}}\right\|_{\mathrm{F}}$ is greater than 2. Observe that the estimation is biaised next to the self-intersection circle of $\mathcal{M}$. Last, let us choose the parameter $\gamma = 2$, and consider the lifted measure $\check{\nu}$. We have $W_1(\check{\mu}_0, \check{\nu}) \approx 0{,}986$. In comparison, the Hausdorff distance between their support is large: $d_{\mathrm{H}}\left(\check{\mathcal{M}}, \operatorname{supp}(\check{\nu})\right) \approx 2{,}188$.

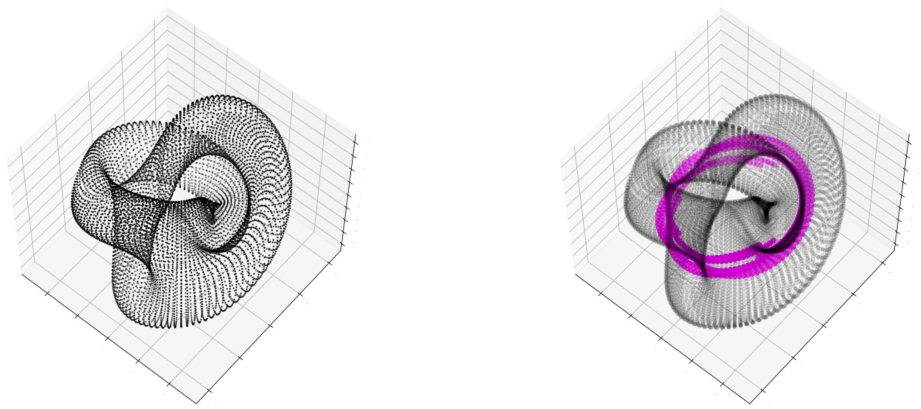

Figure 26: Left: The set X, a sample of $\mathcal{M}$. Right: The set X, where $x \in X$ is colored in magenta if $\left\|\overline{\Sigma}_\nu(x) - \frac{1}{d+2} p_{T_x\mathcal{M}}\right\|_{\mathrm{F}} \geq 2$.

4.2 Homotopy type estimation with the DTM

In this subsection, we use the DTM, as defined in Subsection 1.4, to infer the homotopy type of $\check{\mathcal{M}}$ from the lifted measure $\check{\nu}$. We shall use the DTM on $\check{\nu}$, which lives in the space $E \times \mathcal{M}(E)$ endowed with the norm $\|\cdot\|_\gamma$. It is denoted $d_{\check{\nu},m,\gamma}$.

In order to apply Theorem 1.3 in our setting, we have to consider geometric quantities associated to the submanifold $\check{\mathcal{M}}$. For every $\gamma > 0$, we denote by $\mathrm{reach}_\gamma(\check{\mathcal{M}})$ the reach of $\check{\mathcal{M}}$. Besides, note that the map $\check{u}$ itself satisfies the hypotheses 2 and 3, as the immersion u does. The corresponding constants are denoted $\check{\rho}_\gamma$, $\check{L}_{0,\gamma}$, $\check{f}_{\min,\gamma}$ and $\check{f}_{\max,\gamma}$. We point out that the constant $\check{\rho}_\gamma$ cannot be deduced from ρ: the first one can be arbitrary large or small compared to the second one, even with γ being fixed. This remark holds for the other constants.

However, we can use the results of Section 2 in this context. Proposition 2.18 applied to $\check{\mu}_0$ gives a constant $\check{c}_{5,\gamma}$ such that $\check{\mu}_0(\overline{\mathcal{B}}(\check{x}, r)) \geq \check{c}_{5,\gamma} r^d$ for all $r \leq \frac{1}{2\check{\rho}_\gamma}$. These constants being given, we propose a way to tune the parameters r, γ, m and t in such a way that the t-sublevel set $d^t_{\check{\nu},m,\gamma}$ of the DTM captures the homotopy type of $\check{\mathcal{M}}$, i.e., of $\mathcal{M}_0$.

Corollary 4.3. *Assume that $\mathcal{M}_0$ and μ_0 satisfy the hypotheses 1, 2, 3 and 4. Let ν be any probability measure on E. Denote $w = W_2(\mu, \nu)$. Choose $r > 0$, $\gamma > 0$ and $m \in (0, 1)$ such that*

- $4\left(\frac{w}{c_5 \wedge 1}\right)^{\frac{1}{d+2}} \leq r \leq \frac{1}{2\rho} \wedge r_4 \wedge 1$
- $m \leq \frac{c_{5,\gamma}}{(2\check{\rho}_\gamma)^d}$ *and*
- $(1 + \gamma c_{3.11}) r^{\frac{1}{2}} \leq m^{\frac{1}{2}} \left(\frac{\mathrm{reach}_\gamma(\check{\mathcal{M}})}{9} - \left(\frac{m}{\check{c}_{5,\gamma}}\right)^{\frac{1}{d}}\right)$.

Define ϵ and choose t as follows:

$$\epsilon = \left(\frac{m}{c_{5,\gamma}}\right)^{\frac{1}{d}} + (1 + \gamma c_{3.11}) \left(\frac{r}{m}\right)^{\frac{1}{2}} \qquad \text{and} \qquad t \in \left[4\epsilon, \mathrm{reach}_\gamma(\check{\mathcal{M}}) - 3\epsilon\right].$$

Then the sublevel set of the DTM $d^t_{\check{\nu},m,\gamma}$ is homotopic equivalent to $\mathcal{M}_0$.

Proof. In order to fit in the context of Theorem 1.3, we have to consider the usual Euclidean norm $\|\cdot\|$ on $E \times \mathcal{M}(E)$. It corresponds to the norm $\|\cdot\|_\gamma$ with $\gamma = 1$. For a general parameter $\gamma > 0$, consider the application $i_\gamma \colon E \times \mathcal{M}(E) \to E \times \mathcal{M}(E)$ defined as

$$i_\gamma \colon (x, A) \mapsto (x, \gamma A).$$

A computation shows that, for every probability measures α, β on $E \times \mathcal{M}(E)$, we have

$$W_{2,\gamma}(\alpha, \beta) = W_2\big((i_\gamma)_* \alpha, (i_\gamma)_* \beta\big),$$

where $W_2(\cdot, \cdot)$ denotes the 2-Wasserstein distance on $E \times \mathcal{M}(E)$ endowed with the usual Euclidean norm $\|\cdot\|$. Corollary 3.11 then reformulates as $W_2\big((i_\gamma)_* \check{\mu}_0, (i_\gamma)_* \check{\nu}\big) \leq (1 + \gamma c_{3.11}) r^{\frac{1}{2}}$. Besides, consider the set

$$\check{\mathcal{M}}_\gamma = i_\gamma(\check{\mathcal{M}}) = \{(x, \gamma A), (x, A) \in \check{\mathcal{M}}\}.$$

It is direct to see that

$$\mathrm{reach}_\gamma(\check{\mathcal{M}}) = \mathrm{reach}(\check{\mathcal{M}}_\gamma),$$

where we recall that $\mathrm{reach}_\gamma(\check{\mathcal{M}})$ is the reach of $\check{\mathcal{M}}$ with respect to the norm $\|\cdot\|_\gamma$, and $\mathrm{reach}(\check{\mathcal{M}}_\gamma)$ is the reach of $\check{\mathcal{M}}_\gamma$ with respect to the usual norm $\|\cdot\|$ on $E \times \mathcal{M}(E)$. Finally, consider the DTM $d_{(i_\gamma)_* \check{\nu}, m}$ with respect to the usual Euclidean norm. Observe that, for every $t \geq 0$, the sublevel sets of the DTM $d_{(i_\gamma)_* \check{\nu}, m}$ and $d_{\check{\nu}, m, \gamma}$ are linked via

$$d^t_{\check{\nu}, m} = i_\gamma\left(d^t_{\check{\nu}, m, \gamma}\right).$$

In particular, they share the same homotopy type.

Now we obtain the result as a consequence of Theorem 1.3 applied to $(i_\gamma)_* \check{\mu}_0$ and $(i_\gamma)_* \check{\nu}$. Let us verify that the assumptions of the theorem are satisfied. Our assumption about m ensures that

$$\left(\frac{m}{\check{c}_{5,\gamma}}\right)^{\frac{1}{d}} \leq \frac{1}{2\check{\rho}_\gamma},$$

hence by Proposition 2.18 we get $\check{\mu}_0(\mathcal{B}(x,r)) \geq \check{c}_{5,\gamma} r^d$ for all $x \in \mathrm{supp}(\check{\mu}_0)$ and $r < \left(\frac{m}{\check{c}_{5,\gamma}}\right)^{\frac{1}{d}}$. Moreover, the assumption about $(1+\gamma c_{3.11})r^{\frac{1}{2}}$ ensures that

$$W_2\big((i_\gamma)_*\check{\mu}_0, (i_\gamma)_*\check{\nu}\big) \leq m^{\frac{1}{2}} \left(\frac{\mathrm{reach}_\gamma(\check{\mathcal{M}})}{9} - \left(\frac{m}{c_{5,\gamma}}\right)^{\frac{1}{d}} \right)$$

is satisfied, since $W_2\big((i_\gamma)_*\check{\mu}_0, (i_\gamma)_*\check{\nu}\big) \leq (1+\gamma c_{3.11})r^{\frac{1}{2}}$ by Corollary 3.11. □

Example 4.4. Let $\mathcal{M}$ be the lemniscate of Bernoulli, as in Example 4.1. Suppose that μ is the uniform distribution on $\mathcal{M}$, and ν is the empirical measure on a 500-sample of $\mathcal{M}$. We choose the parameters $\gamma = 2$, $r = 0{,}03$ and $m = 0{,}01$. Let $\check{\nu}$ be the lifted measure associated to ν.

Figure 27 represents set the $\mathrm{supp}(\check{\nu})$, and the values of the DTM $d_{\check{\nu},m,\gamma}$ on it. Observe that the anomalous points, i.e., points for which the local covariance matrix is not well estimated, have large DTM values.

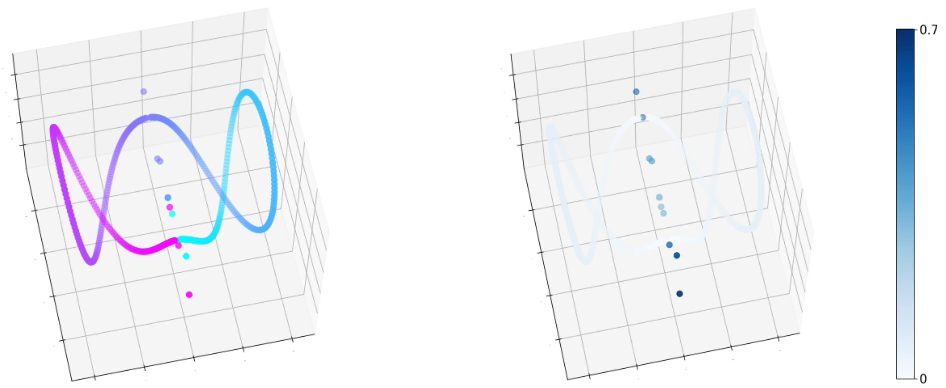

Figure 27: Left: The set $\mathrm{supp}(\check{\nu}) \subset \mathbb{R}^2 \times \mathcal{M}(\mathbb{R}^2)$, projected in a 3-dimensional subspace via PCA. Right: The set $\mathrm{supp}(\check{\nu})$ with colors indicating the value of the DTM $d_{\check{\nu},m,\gamma}$.

4.3 Persistent homology with DTM-filtrations

In this subsection, we aim to estimate the DTM-filtration of $\check{\mu}_0$, as defined in subsection 1.4, from ν. We shall use the DTM-filtration on $\check{\nu}$, denoted $W_\gamma[\check{\nu}]$, with respect to the ambient norm $\|\cdot\|_\gamma$ on $E \times \mathcal{M}(E)$. We use the notations $\check{\rho}_\gamma$ and $c_{5,\gamma}$ of the previous subsection.

Corollary 4.5. *Let $m \in (0,1)$. Assume that $\mathcal{M}_0$ and μ_0 satisfy the hypotheses 1, 2, 3 and 4. Let ν be any probability measure. Denote $W_2(\mu,\nu) = w$. Choose $r > 0$, $\gamma > 0$ and $m \in (0,1)$ such that*

- $4\left(\frac{w}{c_5 \wedge 1}\right)^{\frac{1}{d+2}} \leq r \leq \frac{1}{2\rho} \wedge r_4 \wedge 1,$
- $m \leq \frac{c_{5,\gamma}}{(2\check{\rho}_\gamma)^d},$
- $(1+\gamma c_{3.11})r^{\frac{1}{p}} \leq \frac{1}{4}.$

Then we have a bound on the interleaving distance between the DTM-filtrations:

$$d_i\left(W_\gamma[\check{\mu}_0], W_\gamma[\check{\nu}]\right) \leq \check{c}_{1.6,\gamma}(1+\gamma c_{3.11})^{\frac{1}{2}} m^{-\frac{1}{2}} r^{\frac{1}{4}} + 2\check{c}_{1.4,\gamma} m^{\frac{1}{d}},$$

where $\check{c}_{1.6,\gamma} = 8\mathrm{diam}(\mathcal{M}) + 8\gamma + 5$ and $c_{1.4,\gamma} = (c_{5,\gamma})^{-\frac{1}{d}}$.

Proof. As in the proof of Corollary 4.3, let i_γ be the map $i_\gamma\colon (x, A) \mapsto (x, \gamma A)$. Let $W[\cdot]$ denotes the DTM-filtration on $\check\nu$ with respect to the usual Euclidean norm. That is, the filtration $W[\cdot]$ corresponds to $W_\gamma[\cdot]$ with $\gamma = 1$. A computation shows that the filtration $W[(i_\gamma)_*\check\nu]$ and $W_\gamma[\check\nu]$ are linked via

$$W[(i_\gamma)_*\check\nu] = i_\gamma\left(W_\gamma[\check\nu]\right).$$

Now let $\check w = W_2((i^\gamma)_*\check\mu_0, (i^\gamma)_*\check\nu)$. We have $\check w = W_{2,\gamma}(\check\mu_0, \check\nu)$, hence Corollary 3.11 gives

$$\check w \leq \big(1 + \gamma c_{3.11}\big) r^{\frac{1}{p}}. \tag{31}$$

Moreover, we can apply Corollary 1.6 to $\mu = (i^\gamma)_*\check\mu_0$ and $\nu = (i^\gamma)_*\check\nu$ to get

$$d_i(W[(i^\gamma)_*\check\mu_0], W[(i^\gamma)_*\check\nu]) \leq \check c_{1.6,\gamma}\left(8\mathrm{diam}(\check{\mathcal{M}}) + 5\right)\left(\frac{\check w}{m}\right)^{\frac{1}{2}} + 2\check c_{1.4,\gamma} m^{\frac{1}{d}}, \tag{32}$$

where $\check c_{1.6,\gamma} = \big(8\mathrm{diam}(\check{\mathcal{M}}) + 5\big)$ and $c_{1.4,\gamma} = (c_{5,\gamma})^{-\frac{1}{d}}$. Note that

$$\mathrm{diam}(\check{\mathcal{M}}) \leq \left(\mathrm{diam}(\mathcal{M})^2 + \gamma^2\left(2\frac{1}{2}\right)^2\right)^{\frac{1}{2}} \leq \mathrm{diam}(\mathcal{M}) + \gamma$$

since the matrices $\frac{1}{d+2} p_{T_x\mathcal{M}}$ have norm $\left\|\frac{1}{d+2} p_{T_x\mathcal{M}}\right\|_{\mathrm{F}} = \frac{\sqrt{d}}{d+2} \leq \frac{1}{2}$. Our assumption $m \leq \frac{c_{5,\gamma}}{(2\check\rho_\gamma)^d}$ ensures that the condition $\check\mu_0(\mathcal{B}(x,r)) \geq \check c_{5,\gamma} r^d$ of the Corollary is satisfied. Similarly, the assumption $\big(1 + \gamma c_{3.11}\big) r^{\frac{1}{p}} \leq \frac{1}{4}$ yields $\check w \leq \frac{1}{4}$.

Combining Equations 31 and 32 we get

$$d_i(W[(i^\gamma)_*\check\mu_0], W[(i^\gamma)_*\check\nu]) \leq c_{1.6}\big(1 + \gamma c_{3.11}\big)^{\frac{1}{2}} m^{-\frac{1}{2}} r^{\frac{1}{4}} + 2c_{1.4} m^{\frac{1}{d}}.$$

Now, using the definition of an interleaving of filtrations, one proves that

$$d_i(W_\gamma[\check\mu_0], W_\gamma[\check\nu]) = d_i(W[(i^\gamma)_*\check\mu_0], W[(i^\gamma)_*\check\nu]),$$

and we obtain the result. $\square$

Example 4.6. Say that μ is the uniform measure on the union of five intersecting circles of radius 1. We observe ν, the empirical measure on the point cloud X drawn in Figure 28. It consists in 300 points per circle, and 100 points of clutter noise. Let $p = 1$. Experimentally, we have $W_1(\mu, \nu) \approx 0{,}044$.

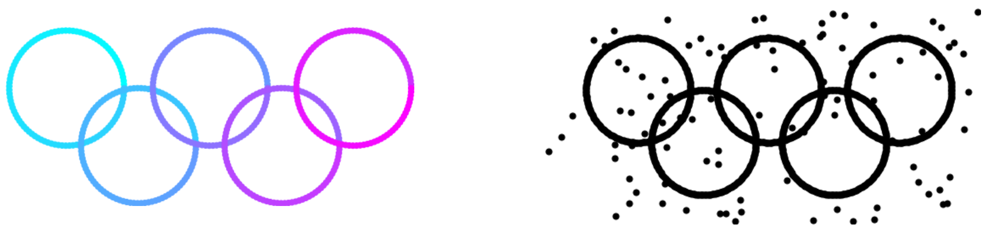

Figure 28: Left: the set $\mathcal{M} = \mathrm{supp}(\mu)$. Right: The set $X = \mathrm{supp}(\nu)$.

Let $\gamma = 1$. Observe that the barcodes of the DTM-filtration $W[(i^\gamma)_*\check\mu_0]$, represented in Figure 29, reveal the homology of the disjoint union of five circles—which is the set $\mathcal{M}_0$. Only bars of length larger than 0,1 are displayed. We consider the construction of $\check\nu$ with parameter $r = 0{,}03$, and the DTM-filtration with $m = 0{,}01$. The barcodes of the DTM-filtration $W[(i^\gamma)_*\check\nu]$ are close to the barcodes of $W[(i^\gamma)_*\check\mu_0]$. To compare, we also plot the persistence diagrams on the usual

Čech filtration on $\mathrm{supp}\,(\check{\nu})$. Observe that the five connected components do not appear clearly anymore.

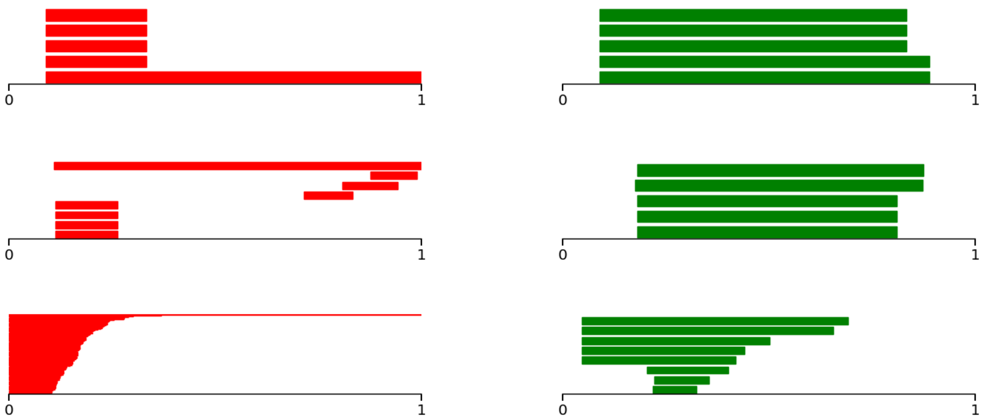

Figure 29: First row: Persistence barcode of the 0- and 1-homology of the DTM-filtration on $\check{\mu}_0$. Second row: Same for $\check{\nu}$. Third row: Persistence barcodes of the usual Čech filtration on $\mathrm{supp}\,(\check{\nu})$.

5 Conclusion

In this paper we described a method to estimate the tangent bundle of a manifold $\mathcal{M}_0$ immersed in the Euclidean space, based on a sample of its image. This estimation is stable in Wasserstein distance. Using the DTM, we are able to estimate the homotopy type of $\mathcal{M}_0$. Moreover, via the DTM-filtrations, we can define a filtration of the space $E \times \mathcal{M}(E)$, whose persistence module contains information about the homology of $\mathcal{M}_0$.

The robust estimation of tangent bundles of manifolds opens the way to the estimation of other topological invariants than homology groups—such as characteristic classes—a problem that will be addressed in a further work.

A Supplementary material for Section 1

Proof of Lemma 1.4. By definition,

$$\delta_{\mu,t}(x) = \inf\left\{r \geq 0,\ \mu\left(\overline{\mathcal{B}}\left(x,r\right)\right) > t\right\} \qquad \text{and} \qquad d^2_{\mu,m}(x) = \frac{1}{m}\int_0^m \delta^2_{\mu,t}(x)dt.$$

Using the assumption $\mu(\mathcal{B}\left(x,r\right)) \geq ar^d$ for all $x \in \mathrm{supp}\,(\mu)$, we get $\delta_{\mu,t}(x) \leq (\frac{t}{a})^{\frac{1}{d}}$, and a simple computation yields

$$d^2_{\mu,m}(x) \leq \frac{d}{d+2}\left(\frac{t}{a}\right)^{\frac{2}{d}} \leq \left(\frac{t}{a}\right)^{\frac{2}{d}}. \qquad \square$$

Proof of Corollary 1.6. Let π be an optimal transport plan for $w = W_2(\mu,\nu)$. Denote $\alpha = w^{\frac{1}{2}}$ and $D = \mathrm{diam}(\mathrm{supp}\,(\mu))$. Define π' to be π restricted to the set $\{x, y \in E, \|x-y\| < \alpha\}$. We denote its marginals μ' and ν'. By Markov inequality, $1 - |\pi'| \leq \frac{w^2}{\alpha^2} = w$. Consider the probability measures $\overline{\mu'}$ and $\overline{\nu'}$. Let us show that we have

$$W_2(\mu,\overline{\mu'}) = 2D\alpha, \qquad W_2(\overline{\mu'},\overline{\nu'}) \leq \alpha \qquad \text{and} \qquad W_2(\nu,\overline{\nu'}) \leq 2(1+D)\alpha. \quad (33)$$

The first inequality is an application of Lemma C.2:

$$W_2(\mu,\overline{\mu'}) \leq 2(1-|\mu'|)^{\frac{1}{2}} D = 2(1-|\pi'|)^{\frac{1}{2}} D \leq 2w^{\frac{1}{2}} D.$$

To obtain the second one, we write

$$W_2^2(\overline{\mu'},\overline{\nu'}) = \int \|x-y\|^2 \,\mathrm{d}\overline{\pi'}(x,y) = \int \|x-y\| \,\frac{\mathrm{d}\pi'(x,y)}{|\pi'|} \leq \frac{1}{|\pi'|}\int \|x-y\|\,\mathrm{d}\pi(x,y).$$

Hence Jensen inequality leads to $W_2(\overline{\mu'},\overline{\nu'}) \leq \frac{w}{|\pi'|^{\frac{1}{2}}}$. Since $1-|\pi'| \leq w$, we have $\frac{w}{|\pi'|^{\frac{1}{2}}} \leq \frac{w}{1-w}$, and the assumption $w \leq \frac{1}{4}$ yields $\frac{w}{1-w} \leq \alpha$. This proves the second point. Finally, we obtain the third inequality by applying the triangular inequality:

$$W_1(\nu,\overline{\nu'}) \leq W_1(\nu,\mu) + W_1(\mu,\overline{\mu'}) + W_1(\overline{\mu'},\overline{\nu'}).$$

Next, let us deduce that

$$c(\overline{\mu'}) \leq c(\mu) + m^{-\frac{1}{2}}2D\alpha \qquad \text{and} \qquad c(\overline{\nu'}) \leq c(\mu) + \left(m^{-\frac{1}{2}} + m^{-\frac{1}{2}}2D + 1\right)\alpha. \tag{34}$$

The first inequality follows from Theorem 1.2:

$$c(\overline{\mu'}) = \sup_{x\in \mathrm{supp}\left(\overline{\mu'}\right)} d_{\overline{\mu'}}(x) \leq \sup_{x\in \mathrm{supp}\left(\overline{\mu'}\right)} d_{\mu}(x) + m^{-\frac{1}{2}} W_2(\overline{\mu'},\mu),$$

and we conclude with $W_2(\mu,\overline{\mu'}) = 2D\alpha$. In order to prove the second inequality, we also use Theorem 1.2:

$$c(\overline{\nu'}) = \sup_{x\in \mathrm{supp}\left(\overline{\nu'}\right)} d_{\overline{\nu'}}(x) \leq \sup_{x\in \mathrm{supp}\left(\overline{\nu'}\right)} d_{\overline{\mu'}}(x) + m^{-\frac{1}{2}} W_2(\overline{\mu'},\overline{\nu'})$$

Since π' has support included in $\{x,y\in E, \|x-y\|<\alpha\}$, we can use Proposition 1.1 to obtain

$$\sup_{x\in \mathrm{supp}\left(\overline{\nu'}\right)} d_{\overline{\mu'}}(x) \leq \sup_{x\in \mathrm{supp}\left(\overline{\mu'}\right)} d_{\overline{\mu'}}(x) + \alpha = c(\mu') + \alpha$$

and we deduce

$$\begin{aligned} c(\overline{\nu'}) &\leq c(\mu') + \alpha + m^{-\frac{1}{2}} W_2(\overline{\mu'},\overline{\nu'}) \\ &\leq c(\mu) + (m^{-\frac{1}{2}} + m^{-\frac{1}{2}}2D + 1)\alpha. \end{aligned}$$

To conclude, Theorem 1.5 gives

$$\begin{aligned} d_i(W[\mu],W[\nu]) &\leq m^{-\frac{1}{2}} W_1(\mu,\overline{\mu'}) + m^{-\frac{1}{2}} W_1(\overline{\mu'},\overline{\nu'}) + m^{-1} W_1(\nu,\overline{\nu'}) + c(\overline{\mu'}) + c(\overline{\nu'}) \\ &\leq \left(m^{-\frac{1}{2}}(4D+1) + 4(D+1)\right)\alpha + 2c(\mu), \end{aligned}$$

where we used Equations 33 and 34 on the last line. Since $m \leq 1$, we can simplify this expression into

$$d_i(W[\mu],W[\nu]) \leq m^{-\frac{1}{2}}(8D+5)\alpha + 2c(\mu).$$

We conclude the proof using $c(\mu) \leq c_{1.4} m^{\frac{1}{d}}$ (Lemma 1.4). $\square$

B Supplementary material for Section 2

Proof of Lemma 2.6. Point (1): We use the triangle inequality, the Pythagorean Theorem and Lemma 2.3 to get

$$\begin{aligned}
\|\gamma(t)-x\| &\geq \|(y+tv)-x\| - \|\gamma(t)-(y+tv)\| \\
&\geq \sqrt{\|tv\|^2 + \|y-x\|^2 + 2\langle tv, y-x\rangle} - \frac{\rho}{2}t^2 \\
&\geq \sqrt{t^2+l^2} - \frac{\rho}{2}t^2.
\end{aligned}$$

Now, a computation shows that the function $t \mapsto \sqrt{t^2+l^2} - \frac{\rho}{2}t^2$ is greater than l on $(0, T_1)$, where $T_1 = \frac{2}{\rho}\sqrt{1-\rho l}$. Hence for $t \in (0, T_1)$, we have $\phi(t) = \|\gamma(t)-x\|^2 > l^2 = \phi(0)$.

Point (2): Observe that $\dot{\phi}(t) = 2\langle \dot{\gamma}(t), \gamma(t)-x\rangle$, and that

$$\ddot{\phi}(t) = 2\langle \dot{\gamma}(t), \dot{\gamma}(t)\rangle + 2\langle \ddot{\gamma}(t), \gamma(t)-x\rangle .$$

By Cauchy-Schwarz inequality, $\langle \ddot{\gamma}(t), \gamma(t)-x\rangle \geq -\|\ddot{\gamma}(t)\|\,\|\gamma(t)-x\|$. Note that $\langle \dot{\gamma}(t), \dot{\gamma}(t)\rangle = 1$ and $\|\ddot{\gamma}(t)\| \leq \rho$. Hence we get

$$\ddot{\phi}(t) \geq 2(1-\rho\|\gamma(t)-x\|). \tag{35}$$

Now, since $\langle v, y-x\rangle = 0$, we have

$$\begin{aligned}
\|\gamma(t)-x\| &\leq \|(y+tv)-x\| + \|\gamma(t)-(y+tv)\| \\
&\leq \sqrt{\|tv\|^2 + \|y-x\|^2} + \frac{\rho}{2}t^2 \\
&= \sqrt{t^2+l^2} + \frac{\rho}{2}t^2.
\end{aligned}$$

A computation shows that the function $t \mapsto \sqrt{t^2+l^2} + \frac{\rho}{2}t^2$ is lower than $\frac{1}{\rho}$ on $(0, T_2)$, where $T_2 = \frac{\sqrt{2}}{\rho}\sqrt{2-\sqrt{3+\rho^2 l^2}}$. Hence for $t \in (0, T_2)$, we have $\ddot{\phi}(t) \geq 0$. And since $\dot{\phi}(0) = 0$, we have that ϕ is increasing.

Point (3): For all $t \in (0, b)$, it holds that $\|\gamma(t)-x\| \leq r$, hence Equation 35 gives $\ddot{\phi}(t) \geq 2(1-\rho r)$.

Point (4): Assume that $\langle v, y-x\rangle \leq 0$. We still have the inequality

$$\|\gamma(t)-x\| \leq \sqrt{t^2+l^2} + \frac{\rho}{2}t^2. \tag{36}$$

Consider t^*, the first non-negative root of $\sqrt{t^2+l^2} + \frac{\rho}{2}t^2 = r$. According to Equation 36, $b \geq t^*$. Now, a computation gives

$$t^* = \frac{\sqrt{2}}{\rho}\sqrt{1+\rho r - \sqrt{(1+\rho r)^2 - \rho^2(r^2-l^2)}}.$$

Using the inequality $\sqrt{B} - \sqrt{A} = \frac{1}{\sqrt{B}+\sqrt{A}}(B-A) \geq \frac{1}{2\sqrt{B}}(B-A)$, where $A < B$, we get

$$1+\rho r - \sqrt{(1+\rho r)^2 - \rho^2(r^2-l^2)} \geq \frac{1}{2(1+\rho r)}\rho^2(r^2-l^2),$$

and we conclude that $t^* \geq \frac{1}{\sqrt{1+\rho r}}\sqrt{r^2-l^2}$.

Point (5): Assume that $\langle v, y-x\rangle \geq 0$ In the same vein as Point 4, we have $\|\gamma(t)-x\| \geq \sqrt{t^2+l^2} - \frac{\rho}{2}t^2$, and we deduce $b \leq t^*$, where t^* is the first positive root of $\sqrt{t^2+l^2} - \frac{\rho}{2}t^2 = r$. Solving this equation leads to

$$t^* = \frac{\sqrt{2}}{\rho}\sqrt{1-\rho r - \sqrt{(1-\rho r)^2 - \rho^2(r^2-l^2)}}.$$

We use the inequality $\sqrt{B}-\sqrt{A} = \frac{1}{\sqrt{A}+\sqrt{B}}(B-A) \le \frac{1}{\sqrt{B}}(B-A)$, where $A < B$, to get

$$1-\rho r - \sqrt{(1-\rho r)^2 - \rho^2(r^2-l^2)} \le \frac{1}{1-\rho r}\rho^2(r^2-l^2)$$

and we conclude that $t^* \le \frac{\sqrt{2}}{\sqrt{1-\rho r}}\sqrt{r^2-l^2}$. $\square$

Proof of Proposition 2.18. Let $\mathcal{M}^x = \mathcal{M} \cap \overline{\mathcal{B}}(x,r)$ and $\mathcal{M}_0^x = u^{-1}(\mathcal{M}^x)$. Lemma 2.10 does not apply: it is not true that $\mathcal{M}_0^x \subseteq \overline{\mathcal{B}}_{\mathcal{M}_0}(x_0, c_{2.10}(\rho r)r)$. However, we can decompose $\mathcal{M}_0^x$ in connected components $C_0^i, i \in I$.

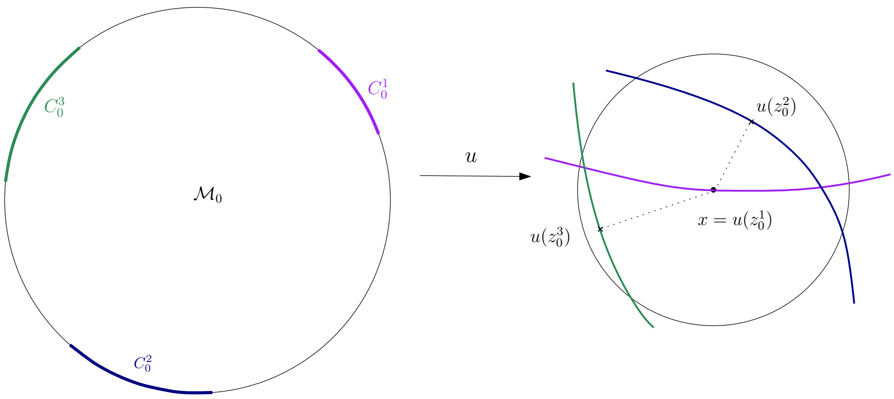

Figure 30: The connected components C_0^i.

For every $i \in I$, let z_0^i be a minimizer of $z_0 \mapsto \|z-x\|$ on C_0^i. We have $x - z^i \perp T_{z^i}\mathcal{M}$, hence according to Lemma 2.6 Point 5, $C_0^i \subseteq \overline{\mathcal{B}}_{\mathcal{M}_0}\left(z_0^i, \frac{1}{\rho}\right)$. For all $i \in I$, consider μ_0^i, the measure μ_0 restricted to C_0^i, and define $\nu_0^i = (\overline{\exp}_{z_0}^{\mathcal{M}_0})_*^{-1}\mu_0^i$, as in Remark 2.16. The measure ν_0^i admits g_0^i as a density over the d-dimensional Hausdorff measure on $T_{z_0^i}\mathcal{M}_0$, where

$$g_0^i(v) = f_0\big(\overline{\exp}_{z_0}^{\mathcal{M}_0}(v)\big) \cdot J_v \cdot 1_{(\overline{\exp}_{z_0}^{\mathcal{M}_0})^{-1}(C_0^i)}(v).$$

Point (1): We can write

$$\mu(\overline{\mathcal{B}}(x,r)) = \mu_0(u^{-1}(\overline{\mathcal{B}}(x,r))) = \sum_{i \in I} \mu_0(C_0^i).$$

Let $i_* \in I$ be the index of the connected component of $\mathcal{M}_0^x$ which contains x_0. We have $C_0^{i_*} \supset \overline{\mathcal{B}}_{\mathcal{M}_0}(x_0, r)$, and we deduce that

$$\begin{aligned}\mu_0(C_0^{i_*}) &\ge \int_{(\overline{\exp}_{z_0}^{\mathcal{M}_0})^{-1}(C_0^i)} g_0^{i_*}\,\mathrm{d}\mathcal{H}^d \\ &\ge f_{\min} J_{\min} \mathcal{H}^d\big((\overline{\exp}_{z_0}^{\mathcal{M}_0})^{-1}(C_0^i)\big) = f_{\min} J_{\min} V_d r^d.\end{aligned}$$

Therefore, $\mu(\overline{\mathcal{B}}(x,r)) \ge f_{\min} J_{\min} V_d r^d$.

Point (2): We now prove the second point.

Step 1: Let us show that the cardinal of I is lower than $\frac{1}{f_{\min} J_{\min} V_d}(\frac{2\rho}{\alpha})^d$, with $\alpha = \sqrt{4-\sqrt{13}}$. We shall prove that for every $i, j \in I$ such that $i \ne j$, $\mathrm{d}_{\mathcal{M}_0}\left(z_0^i, z_0^j\right) \ge \frac{\alpha}{\rho}$.

Let γ_0 be a geodesic from z_0^i to z_0^j, with $\gamma(0) = z^i$, $\gamma(T) = z^j$, and $\dot\gamma_0(0) = v_0$. Consider the application $\phi\colon t \mapsto \|\gamma(t) - x\|^2$. Since C_0^i and C_0^j are disjoint connected components, there must be a $t^* < T$ such that $\|\gamma(t^*) - x_0\| > r$. Moreover, according to Lemma 2.6 Point 2, ϕ is increasing on $[0, T_2]$ where $T_2 = \frac{\sqrt{2}}{\rho}\sqrt{2 - \sqrt{3 + \rho^2 l^2}}$. Since $\phi(T) \le r$, we deduce that T is greater than T_2. Note that the assumption $r \le \frac{1}{2\rho}$ yields $T_2 \ge \frac{\alpha}{\rho}$.

This implies that the geodesic balls $\mathcal{B}_{\mathcal{M}_0}\left(z_0^i, \frac{\alpha}{2\rho}\right)$ are disjoint. Therefore,

$$1 \ge \mu_0\big(\bigcup_i \mathcal{B}_{\mathcal{M}_0}\left(z_0^i, \frac{\alpha}{2\rho}\right)\big) \ge |I| f_{\min} J_{\min} V_d (\frac{\alpha}{2\rho})^d,$$

and we deduce $|I| \le \frac{1}{f_{\min} J_{\min} V_d} (\frac{2\rho}{\alpha})^d$.

Step 2: Let $i \in I$, and define $D_0^i = C_0^i \cap u^{-1}(\overline{\mathcal{B}}(x, r) \setminus \overline{\mathcal{B}}(x, s))$. Let us show that

$$\mu_0(D_0^i) \le f_{\max} J_{\max} 2^{d-1} \sqrt{6} d V_d \cdot r^{d-1} \sqrt{r^2 - s^2}.$$

Let us distinguish two cases: $l_i \ge s$ or $l_i < s$.

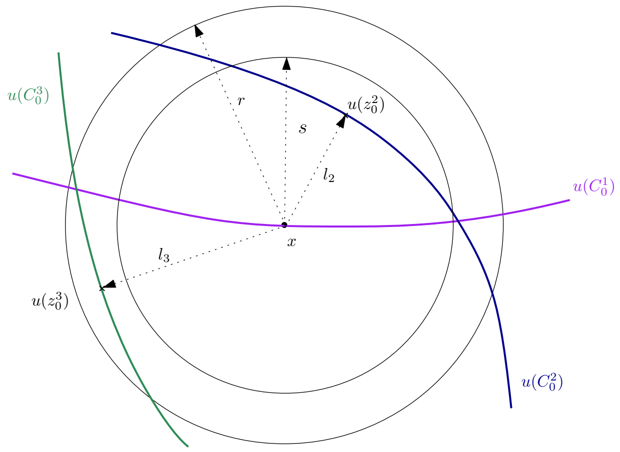

Figure 31: Illustration of the cases $l_i \ge s$ and $l_i < s$.

First, assume that $l_i < s$. Let γ be a geodesic starting from z_0^i, denote $v = \dot{\gamma}(0)$ and consider the application $\phi \colon t \mapsto \|\gamma(t) - x\|^2$. Let $a(v), b(v)$ be the first values of $t \ge 0$ such that $\|\gamma(t) - x\| = s$ and $\|\gamma(t) - x\| = r$. As in the proof of Proposition 2.17 Point 3, we still have Equation 7:

$$r^2 - s^2 \ge (1 - \rho r)(b(v)^2 - a(v)^2),$$

from which we deduce $b(v) - a(v) \le \frac{1}{1-\rho r} \frac{1}{b(v)+a(v)} (r^2 - s^2)$. According to Lemma 2.6 Point 4, $b(v) + a(v) \ge b(v) \ge (1 + \rho r)^{-\frac{1}{2}} \sqrt{r^2 - l_i^2} \ge (1 + \rho r)^{-\frac{1}{2}} \sqrt{r^2 - s^2}$, and we obtain

$$b(v) - a(v) \le \frac{(1 + \rho r)^{\frac{1}{2}}}{1 - \rho r} \sqrt{r^2 - s^2}. \tag{37}$$

Now, we write

$$\mu_0(D_0^i) = \nu_0^i\big((\overline{\exp}_{z_0^i}^{\mathcal{M}_0})^{-1}(D_0^i)\big).$$

In spherical coordinates, this measure reads

$$\int_{(\overline{\exp}_{z_0^i}^{\mathcal{M}_0})^{-1}(D_0^i)} g_0^i(y) \mathrm{d}\mathcal{H}^d(y) = \int_{v \in \partial\mathcal{B}(0,1)} \int_{t=a(v)}^{b(v)} g_0^i(tv) t^{d-1} \mathrm{d}t \mathrm{d}v.$$

We can now conclude as in the proof of Proposition 2.17 Point 3. We still have $b(v) \leq 2r$, and we write

$$\int_{t=a(v)}^{b(v)} g_0^i(tv)t^{d-1}\mathrm{d}t \leq \int_{t=a(v)}^{b(v)} f_{\max} J_{\max}(2r)^{d-1}\mathrm{d}t.$$

Using Equation 37, we obtain

$$\int_{t=a(v)}^{b(v)} f_{\max} J_{\max}(2r)^{d-1}\mathrm{d}t \leq \frac{(1+\rho r)^{\frac{1}{2}}}{1-\rho r}\sqrt{r^2-s^2} f_{\max} J_{\max}(2r)^{d-1}.$$

Therefore,

$$\int_{v\in\partial\mathcal{B}(0,1)}\int_{t=a(v)}^{b(v)} g_0^i(tv)t^{d-1}\mathrm{d}t\mathrm{d}v \leq \frac{(1+\rho r)^{\frac{1}{2}}}{1-\rho r}\sqrt{r^2-s^2} f_{\max} J_{\max}(2r)^{d-1} dV_d.$$

The assumption $r < \frac{1}{2\rho}$ yields $\frac{(1+\rho r)^{\frac{1}{2}}}{1-\rho r} < \sqrt{6}$, and we finally obtain

$$\mu_0(D_0^i) \leq f_{\max} J_{\max} 2^{d-1}\sqrt{6} dV_d \cdot r^{d-1}\sqrt{r^2-s^2}.$$

Now, assume that $l_i \geq s$. This case is similar to the first one. One has

$$\mu_0(D_0^i) \leq \int_{(\overline{\exp}_{z_0^i}^{\mathcal{M}_0})^{-1}(D_0^i)} g_0^i(y)\mathrm{d}\mathcal{H}^d(y) = \int_{v\in\partial\mathcal{B}(0,1)}\int_{t=0}^{b(v)} g_0(tv)t^{d-1}\mathrm{d}t\mathrm{d}v.$$

and Lemma 2.6 Point 5 gives $b(v) \leq (\frac{1-\rho r}{2})^{-\frac{1}{2}}\sqrt{r^2-l^2} \leq (\frac{1-\rho r}{2})^{-\frac{1}{2}}\sqrt{r^2-s^2}$. Note that $(\frac{1-\rho r}{2})^{-\frac{1}{2}}$ is not greater than 2 when $r < \frac{1}{2\rho}$. One deduces that

$$\mu_0(D_0^i) \leq f_{\max} J_{\max} 2^{d-1} 2 dV_d \cdot r^{d-1}\sqrt{r^2-s^2}.$$

Step 3: We conclude: since $u^{-1}(\overline{\mathcal{B}}(x,r) \setminus \overline{\mathcal{B}}(x,s)) = \bigcup_i D_0^i$, Step 1 and 2 yield

$$\begin{aligned}\mu(\overline{\mathcal{B}}(x,r) \setminus \overline{\mathcal{B}}(x,s)) = \sum_{i\in I}\mu_0(D_i) &\leq |I| f_{\max} J_{\max} 2^{d-1}\sqrt{6} dV_d \cdot r^{d-1}\sqrt{r^2-s^2} \\ &\leq \frac{1}{f_{\min} J_{\min} V_d}(\frac{2\rho}{\alpha})^d f_{\max} J_{\max} 2^{d-1}\sqrt{6} dV_d \cdot r^{d-1}\sqrt{r^2-s^2}.\end{aligned}$$

Finally, the inequality $\sqrt{r^2-s^2} \leq \sqrt{2r}\sqrt{r-s}$ yields

$$\mu(\overline{\mathcal{B}}(x,r) \setminus \overline{\mathcal{B}}(x,s)) \leq \frac{f_{\max} J_{\max}}{f_{\min} J_{\min}}(\frac{\rho}{\alpha})^d d 2^{2d}\sqrt{3} r^{d-\frac{1}{2}}\sqrt{r-s}. \qquad \square$$

C Supplementary material for Section 3

In this subsection, we suppose that μ and ν are probability measures on E.

Lemma C.1. *For every $x, y \in E$, we have $\left\|x^{\otimes 2} - y^{\otimes 2}\right\|_{\mathrm{F}} \leq (\|x\| + \|y\|)\|x-y\|$.*

Proof. We apply the triangular inequality to $x^t x - y^t y = (x-y)^t x + y^t(x-y)$:

$$\begin{aligned}\left\|x^t x - y^t y\right\|_{\mathrm{F}} \leq \left\|(x-y)^t x\right\|_{\mathrm{F}} + \left\|y^t(x-y)\right\|_{\mathrm{F}} &\leq \|x-y\|\|x\| + \|y\|\|x-y\| \\ &= (\|x\|+\|y\|)\|x-y\|. \qquad \square\end{aligned}$$

Lemma C.2. *Let μ' be a submeasure of μ with $|\mu'| > 0$, and consider the corresponding probability measure $\overline{\mu'}$. Suppose that* $\operatorname{supp}(\mu)$ *is included in a ball $\overline{\mathcal{B}}(x, r)$. Then*

$$W_p\left(\mu, \overline{\mu'}\right) \leq 2(1 - |\mu'|)^{\frac{1}{p}} r.$$

More generally, let μ be any measure of positive mass (potentially with $|\mu| \neq 1$), and let μ' be a submeasure of μ with $|\mu'| > 0$. Suppose that $\operatorname{supp}(\mu)$ *is included in a ball $\overline{\mathcal{B}}(x, r)$. Then*

$$W_p\left(\overline{\mu}, \overline{\mu'}\right) \leq 2\left(1 - \frac{|\mu'|}{|\mu|}\right)^{\frac{1}{p}} r.$$

Proof. We start with the first inequality. Consider the intermediate probability measure $\omega = \mu' + (1 - |\mu'|)\delta_x$. We shall use the triangular inequality $W_1(\mu, \overline{\mu'}) \leq W_1(\mu, \omega) + W_1(\omega, \overline{\mu'})$. We can write

- $\mu = \mu' + (\mu - \mu')$,
- $\omega = \mu' + (1 - |\mu'|)\delta_x$,
- $\overline{\mu'} = \mu' + (\overline{\mu'} - \mu')$.

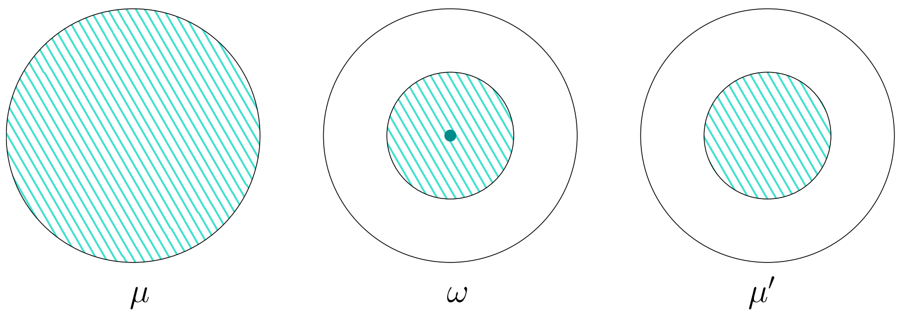

Figure 32: The measures involved in the proof of Lemma C.2. A hatched area represents the support of the measure, and a point represents a Dirac mass.

Observe that μ and ω admits μ' as a common submeasure of mass $|\mu'|$. Therefore we can build a transport plan between μ and ω where only a mass $1 - |\mu'|$ of μ is moved to x. In other words,

$$W_p(\mu, \omega) \leq (1 - |\mu'|)^{\frac{1}{p}} r.$$

Similarly, one shows that $W_p\left(\omega, \overline{\mu'}\right) \leq (1 - |\mu'|)^{\frac{1}{p}} r$.

Now let us prove the second inequality. Since μ' is a submeasure of μ of mass $|\mu'|$, then $\frac{1}{|\mu|}\mu'$ is a submeasure of $\overline{\mu} = \frac{1}{|\mu|}\mu$ of mass $\frac{1}{|\mu|}|\mu'|$. We then apply the first inequality. □

Lemma C.3. *Let $x \in \operatorname{supp}(\mu)$. Suppose that x satisfies the Hypotheses 5 and 6 with $\lambda(x) \wedge \frac{1}{2\rho} > r$. Let $y \in E$ such that $\|x - y\| < \frac{r}{4}$. Then $|\mu_x|, |\mu_y| > 0$, and*

$$W_1\left(\overline{\mu_x}, \overline{\mu_y}\right) \leq c_{\mathrm{C.3}} \|x - y\|$$

with $c_{\mathrm{C.3}} = 2\left(1 + 4\frac{5^{d-1}}{3^d}\right)\frac{c_6}{c_5}$.

Proof. It is clear that $|\mu_y| > 0$ since $\mu(\overline{\mathcal{B}}(y, r)) \geq \mu(\overline{\mathcal{B}}(x, r - \|x - y\|))$ and $x \in \operatorname{supp}(\mu)$. Let us show the inequality $W_1(\overline{\mu_x}, \overline{\mu_y}) \leq c_{\mathrm{C.3}} \|x - y\|$ by studying the measure μ on the intersection

$\overline{\mathcal{B}}(x,r)\cap\overline{\mathcal{B}}(y,r)$. Let $\mu_{x,y}$ be the restriction of μ to $\overline{\mathcal{B}}(x,r)\cap\overline{\mathcal{B}}(y,r)$, and $\overline{\mu_{x,y}}$ the corresponding probability measure. The triangular inequality gives:

$$W_1(\overline{\mu_x},\overline{\mu_y}) \leq \underbrace{W_1(\overline{\mu_x},\overline{\mu_{x,y}})}_{(1)} + \underbrace{W_1(\overline{\mu_{x,y}},\overline{\mu_y})}_{(2)}.$$

Term (1): Let us show that $W_1(\overline{\mu_x},\overline{\mu_{x,y}}) \leq 2\frac{c_6}{c_5}\|x-y\|$. Note that $\overline{\mu_{x,y}}$ is a submeasure of $\overline{\mu_x}$. According to Lemma C.2, we have

$$W_1(\overline{\mu_x},\overline{\mu_{x,y}}) \leq 2\left(1-\frac{|\mu_{x,y}|}{|\mu_x|}\right)r = 2\frac{|\mu_x|-|\mu_{x,y}|}{|\mu_x|}r.$$

We know from Hypothesis 5 that $|\mu_x| \geq c_5 r^d$. On the other hand,

$$\begin{aligned}|\mu_x|-|\mu_{x,y}| &= \mu(\overline{\mathcal{B}}(x,r)) - \mu(\overline{\mathcal{B}}(x,r)\cap\overline{\mathcal{B}}(y,r))\\ &\leq \mu(\overline{\mathcal{B}}(x,r)) - \mu(\overline{\mathcal{B}}(x,r-\|x-y\|)),\end{aligned}$$

hence we can apply Hypothesis 6 to get $|\mu_x|-|\mu_{x,y}| \leq c_6 r^{d-1}\|x-y\|$. We finally obtain

$$W_1(\overline{\mu_x},\overline{\mu_{x,y}}) \leq 2\frac{c_6 r^{d-1}\|x-y\|}{c_5 r^d}r = 2\frac{c_6}{c_5}\|x-y\|.$$

Term (2): Similarly, Lemma C.2 yields

$$W_1(\overline{\mu_y},\overline{\mu_{x,y}}) \leq 2\frac{|\mu_y|-|\mu_{x,y}|}{|\mu_y|}r.$$

Let us show that we still have $|\mu_y| \geq a'r^d$ and $|\mu_y|-|\mu_{x,y}| \leq b'r^{d-1}\|x-y\|$ with the constants $a' = (\frac{3}{4})^d c_5$ and $b' = 2(\frac{5}{4})^{d-1}c_6$. The first inequality comes from Hypothesis 5:

$$\mu(\overline{\mathcal{B}}(y,r)) \geq \mu(\overline{\mathcal{B}}(x,r-\|x-y\|)) \geq c_5(r-\|x-y\|)^d$$

and $\|x-y\| \leq \frac{r}{4}$. The second inequality comes from Hypothesis 6:

$$\begin{aligned}\mu(\overline{\mathcal{B}}(y,r)) - \mu(\overline{\mathcal{B}}(x,r)\cap\overline{\mathcal{B}}(y,r)) &\leq \mu(\overline{\mathcal{B}}(x,r+\|x-y\|)) - \mu(\overline{\mathcal{B}}(x,r-\|x-y\|))\\ &\leq c_6(r+\|x-y\|)^{d-1}2\|x-y\|\end{aligned}$$

and $\|x-y\| \leq \frac{r}{4}$. To conclude,

$$W_1(\overline{\mu_y},\overline{\mu_{x,y}}) \leq 2\frac{2(\frac{5}{4})^{d-1}r^{d-1}c_5\|x-y\|}{2(\frac{3}{4})^d c_6 r^d}r = 8\frac{5^{d-1}}{3^d}\frac{c_6}{c_5}\|x-y\|. \qquad \square$$

Lemma C.4. *Let $x \in \operatorname{supp}(\mu)$. Suppose that x satisfies the Hypotheses 5 and 7 at x with $\frac{1}{2\rho} > r$. Let $y \in E$ such that $\|x-y\| < \frac{r}{4}$. Then $|\mu_x|, |\mu_y| > 0$, and*

$$W_1(\overline{\mu_x},\overline{\mu_y}) \leq c_{\mathrm{C.4}} r^{\frac{1}{2}}\|x-y\|^{\frac{1}{2}}$$

with $c_{\mathrm{C.4}} = \left(2+\frac{2^{\frac{5}{2}}5^{d-\frac{1}{2}}}{3^d}\right)\frac{c_7}{c_5}$.

Proof. The proof is similar to Lemma C.3 with slight modifications. We still consider

$$W_1(\overline{\mu_x},\overline{\mu_y}) \leq \underbrace{W_1(\overline{\mu_x},\overline{\mu_{x,y}})}_{(1)} + \underbrace{W_1(\overline{\mu_{x,y}},\overline{\mu_y})}_{(2)}.$$

Term (1): We have $W_1(\overline{\mu_x},\overline{\mu_{x,y}}) \leq 2\frac{|\mu_x|-|\mu_{x,y}|}{|\mu_x|}r$. Hypothesis 5 still gives $|\mu_x| \geq c_5 r^d$. But Hypothesis 7 now yields

$$\begin{aligned}|\mu_x|-|\mu_{x,y}| &\leq \mu(\overline{\mathcal{B}}(x,r)) - \mu(\overline{\mathcal{B}}(x,r-\|x-y\|))\\ &\leq c_7 r^{d-\frac{1}{2}}\|x-y\|^{\frac{1}{2}}.\end{aligned}$$

We finally obtain $W_1(\overline{\mu_x},\overline{\mu_{x,y}}) \leq 2\frac{c_7}{c_5}r^{\frac{1}{2}}\|x-y\|^{\frac{1}{2}}$.

Term (2): In order to bound $W_1(\overline{\mu_y}, \overline{\mu_{x,y}}) \leq 2\frac{|\mu_y|-|\mu_{x,y}|}{|\mu_y|}r$, Hypothesis 5 still gives $|\mu_x| \geq (\frac{3}{4})^d c_5 r^d$, and Hypothesis 7 yields

$$\begin{aligned}|\mu_y| - |\mu_{x,y}| &\leq \mu(\overline{\mathcal{B}}(x, r + \|x-y\|)) - \mu(\overline{\mathcal{B}}(x, r - \|x-y\|)) \\ &\leq c_7(r + \|x-y\|)^{d-\frac{1}{2}}(2\|x-y\|)^{\frac{1}{2}},\end{aligned}$$

which is not greater than $c_7(\frac{5}{4}r)^{d-\frac{1}{2}}(2\|x-y\|)^{\frac{1}{2}}$.

We finally get $W_1(\overline{\mu_y}, \overline{\mu_{x,y}}) \leq 2\frac{c_7(\frac{5}{4}r)^{d-\frac{1}{2}}(2\|x-y\|)^{\frac{1}{2}}}{(\frac{3}{4})^d c_5 r^d}r \leq \frac{2^{\frac{5}{2}}5^{d-\frac{1}{2}}c_7}{3^d c_5}r^{\frac{1}{2}}\|x-y\|^{\frac{1}{2}}$. □

Lemma C.5. *Let $w = W_p(\mu,\nu)$. Let $y \in E$. Suppose that there exists $x \in \operatorname{supp}(\mu)$ such that $\|x-y\| \leq \alpha$ with $\alpha = (\frac{w}{r^{d-1}})^{\frac{1}{2}}$, and that μ satisfies the Hypotheses 5 and 6 at x with $\lambda(x) \wedge \frac{1}{2\rho} > r$. Assume that $w \leq (c_5 \wedge 1)(\frac{r}{4})^{d+1}$. Then*

$$W_1(\overline{\mu_y}, \overline{\nu_y}) \leq c_{C.5}\alpha$$

with $c_{C.5} = \frac{2^{d-1}}{c_5} + 2\frac{12\cdot 5^{d-1}c_6+1}{3^d c_5} + 2^{d+3}\frac{(\frac{3}{2})^{d-1}c_6+1}{c_5}$.

Proof. Let π be an optimal transport for $W_p(\mu,\nu)$. Define π_y to be the restriction of the measure π to the set $\overline{\mathcal{B}}(y,r) \times \overline{\mathcal{B}}(y,r) \subset E \times E$. Its marginals $p_{1*}\pi_y$ and $p_{2*}\pi_y$ are submeasures of μ_y and ν_y. We shall use the triangular inequality:

$$W_1(\overline{\mu_y}, \overline{\nu_y}) \leq \underbrace{W_1(\overline{\mu_y}, \overline{p_{1*}\pi_y})}_{(1)} + \underbrace{W_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y})}_{(2)} + \underbrace{W_1(\overline{p_{2*}\pi_y}, \overline{\nu_y})}_{(3)}$$

Before examinating each of these terms, note that we have

$$|\pi_y| = |p_{1*}\pi_y| = |p_{2*}\pi_y| \geq \mu(\overline{\mathcal{B}}(y, r-\alpha)) - \frac{w}{\alpha} \tag{38}$$

$$|\nu_y| \leq \mu(\overline{\mathcal{B}}(y, r+\alpha)) + \frac{w}{\alpha} \tag{39}$$

$$|\nu_y| \geq \mu(\overline{\mathcal{B}}(y, r-\alpha)) - \frac{w}{\alpha} \tag{40}$$

The first equation can be proven as follows:

$$\begin{aligned}\mu(\overline{\mathcal{B}}(y, r-\alpha)) &= \pi(\overline{\mathcal{B}}(y, r-\alpha) \times E) \\ &= \pi(\overline{\mathcal{B}}(y, r-\alpha) \times \overline{\mathcal{B}}(y,r)) + \pi(\overline{\mathcal{B}}(y, r-\alpha) \times \overline{\mathcal{B}}(y,r)^c)\end{aligned}$$

On the one hand, $\pi(\overline{\mathcal{B}}(y, r-\alpha) \times \overline{\mathcal{B}}(y,r)) \leq \pi(\overline{\mathcal{B}}(y,r) \times \overline{\mathcal{B}}(y,r)) \leq |\pi_y|$. On the other hand, Markov inequality yields

$$\pi(\overline{\mathcal{B}}(y, r-\alpha) \times \overline{\mathcal{B}}(y,r)^c) \leq \pi(\{(z,z'), \|z-z'\| \geq \alpha\}) \leq \frac{1}{\alpha}\int \|z-z'\| \,\mathrm{d}\pi(z,z'),$$

and Jensen inequality gives

$$\frac{1}{\alpha}\int \|z-z'\| \,\mathrm{d}\pi(z,z') \leq \frac{1}{\alpha}\left(\int \|z-z'\|^p \,\mathrm{d}\pi(z,z')\right)^{\frac{1}{p}} = \frac{w}{\alpha}.$$

We deduce that $\mu(\overline{\mathcal{B}}(y, r-\alpha)) \leq |\pi_y| + \frac{w}{\alpha}$, which gives Equation 38. Equations 39 and 40 can be proven similarly.

In addition, note that the assumption $w \leq (c_5 \wedge 1)(\frac{r}{4})^{d+1}$ yields

$$\alpha \leq \frac{r}{4} \tag{41}$$

$$\frac{w}{\alpha} \leq \frac{c_5}{2}\left(\frac{r}{2}\right)^d \tag{42}$$

We now study the terms (1), (2) and (3).

Term (2): Since $\overline{\pi_y} = \frac{\pi_y}{|\pi_y|}$ is a transport plan between $\overline{p_{1*}\pi_y}$ and $\overline{p_{2*}\pi_y}$, we have

$$W_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y}) \leq \int \|z - z'\| \frac{\mathrm{d}\pi_y(z,z')}{|\pi_y|} \leq \frac{1}{|\pi_y|} \int \|z - z'\| \, \mathrm{d}\pi(z,z').$$

Moreover, Jensen inequality yields $\int \|z - z'\| \, \mathrm{d}\pi(z,z') \leq w$. Hence

$$W_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y}) \leq \frac{w}{|\pi_y|}.$$

Let us prove that $|\pi_y| \geq \frac{c_5}{2}(\frac{r}{2})^d$. According to Equation 38, $|\pi_y| \geq \mu(\overline{\mathcal{B}}(y, r-\alpha)) - \frac{w}{\alpha}$. Now, remark that $\mu(\overline{\mathcal{B}}(y, r-\alpha)) \geq \frac{c_5}{2^d} r^d$. Indeed, using Hypothesis 5,

$$\mu(\overline{\mathcal{B}}(y, r-\alpha)) \geq \mu(\overline{\mathcal{B}}(x, r-\alpha - \|x-y\|)) \geq c_5(r - \alpha - \|x-y\|)^d,$$

and we conclude with $\|x-y\| \leq \alpha \leq \frac{r}{4}$. Now, using Equation 42, we get

$$\begin{aligned} |\pi_y| &\geq \mu(\overline{\mathcal{B}}(y, r-\alpha)) - \frac{w}{\alpha} \\ &\geq c_5 \left(\frac{r}{2}\right)^d - \frac{c_5}{2}\left(\frac{r}{2}\right)^d \geq \frac{c_5}{2}\left(\frac{r}{2}\right)^d \end{aligned}$$

Finally, since $\alpha = \left(\frac{w}{r^{d-1}}\right)^{\frac{1}{2}}$ and $\alpha \leq \frac{r}{4}$, we obtain

$$W_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y}) \leq \frac{w}{|\pi_y|} \leq \frac{w}{\frac{c_5}{2}(\frac{r}{2})^d} = \frac{2^{d+1}}{c_5}\alpha^2 \frac{1}{r} \leq \frac{2^{d-1}}{c_5}\alpha.$$

Term (1): According to Lemma C.2, we have

$$W_1(\overline{\mu_y}, \overline{p_{1*}\pi_y}) \leq 2\frac{|\mu_y| - |p_{1*}\pi_y|}{|\mu_y|} r.$$

We can use Equation 38 to get

$$\begin{aligned} |\mu_y| - |p_{1*}\pi_y| &\leq \mu(\overline{\mathcal{B}}(y, r)) - \mu(\overline{\mathcal{B}}(y, r-\alpha)) + \frac{w}{\alpha} \\ &\leq \mu(\overline{\mathcal{B}}(x, r + \|x-y\|)) - \mu(\overline{\mathcal{B}}(x, r - \alpha - \|x-y\|)) + \frac{w}{\alpha}. \end{aligned}$$

By Hypothesis 6,

$$\mu(\overline{\mathcal{B}}(x, r + \|x-y\|)) - \mu(\overline{\mathcal{B}}(x, r - \alpha - \|x-y\|)) \leq c_6(r + \|x-y\|)^{d-1}(2\|x-y\| + \alpha),$$

which is not greater than $c_6(\frac{5}{4}r)^{d-1}3\alpha$ since $\|x-y\| \leq \alpha \leq \frac{r}{4}$. Moreover, $\frac{w}{\alpha} = r^{d-1}\alpha$, and we obtain

$$|\mu_y| - |p_{1*}\pi_y| \leq \left(3\left(\frac{5}{4}\right)^{d-1} c_6 + 1\right) r^{d-1}\alpha,$$

Finally, thanks to Hypothesis 5, we write

$$\begin{aligned} |\mu_y| = \mu(\overline{\mathcal{B}}(y, r)) &\geq \mu(\overline{\mathcal{B}}(x, r - \|x-y\|)) \\ &\geq c_5(r - \|x-y\|)^d \geq c_5\left(\frac{3}{4}\right)^d r^d \end{aligned}$$

and we obtain

$$\frac{|\mu_y| - |p_{1*}\pi_y|}{|\mu_y|} \leq \frac{((3(\frac{5}{4})^{d-1}c_6 + 1)r^{d-1}}{c_5(\frac{3}{4})^d r^d}\alpha = \frac{1}{r} \cdot \frac{12 \cdot 5^{d-1}c_6 + 1}{3^d c_5}\alpha.$$

We deduce

$$W_1(\overline{\mu_y}, \overline{p_{1*}\pi_y}) \leq 2\frac{12 \cdot 5^{d-1} c_6 + 1}{3^d c_5} \alpha.$$

Term (3): It is similar to Term (1). First, one shows that

$$W_1(\overline{\nu_y}, \overline{p_{2*}\pi_y}) \leq 2\frac{|\nu_y| - |p_{2*}\pi_y|}{|\nu_y|} r.$$

Using Equations 38 and 39 we get

$$\begin{aligned}|\nu_y| - |p_{2*}\pi_y| &\leq \mu(\overline{\mathcal{B}}(y, r+\alpha)) + \frac{w}{\alpha} - \mu(\overline{\mathcal{B}}(y, r-\alpha)) + \frac{w}{\alpha} \\ &\leq \mu(\overline{\mathcal{B}}(x, r + \|x-y\| + \alpha)) - \mu(\overline{\mathcal{B}}(x, r - \alpha - \|x-y\|)) + 2\frac{w}{\alpha}.\end{aligned}$$

By Hypothesis 6, we have

$$\begin{aligned}&\mu(\overline{\mathcal{B}}(x, r + \|x-y\| + \alpha)) - \mu(\overline{\mathcal{B}}(x, r - \alpha - \|x-y\|)) \\ &\leq c_6(r + \|x-y\| + \alpha)^{d-1}(2\|x-y\| + 2\alpha)\end{aligned}$$

which is not greater than $c_6(\frac{3}{2}r)^{d-1}4\alpha$ since $\|x-y\| \leq \alpha \leq \frac{r}{4}$. Moreover, $\frac{w}{\alpha} = r^{d-1}\alpha$, and we obtain

$$|\nu_y| - |p_{2*}\pi_y| \leq (4(\frac{3}{2})^{d-1} c_6 + 2) r^{d-1}\alpha.$$

We have seen that

$$|\nu_y| \geq \mu(\overline{\mathcal{B}}(y, r-\alpha)) - \frac{w}{\alpha} \geq \frac{c_5}{2}\left(\frac{r}{2}\right)^d.$$

Hence

$$\frac{|\nu_y| - |p_{2*}\pi_y|}{|\nu_y|} \leq \frac{(4(\frac{3}{2})^{d-1} c_6 + 2) r^{d-1}}{\frac{c_5}{2}(\frac{r}{2})^d}\alpha = \frac{1}{r} \cdot 2^{d+2}\frac{(\frac{3}{2})^{d-1} b + 1}{c_5}\alpha,$$

and we finally obtain

$$W_1(\overline{\mu_y}, \overline{p_{1*}\pi_y}) \leq 2^{d+3}\frac{(\frac{3}{2})^{d-1} c_6 + 1}{c_5}\alpha.$$

To conclude, summing up these three terms gives $W_1(\overline{\mu_y}, \overline{\nu_y}) \leq c_{C.5}\alpha$ with $c_{C.5} = \frac{2^{d-1}}{c_5} + 2\frac{12 \cdot 5^{d-1} c_6 + 1}{3^d c_5} + 2^{d+3}\frac{(\frac{3}{2})^{d-1} c_6 + 1}{c_5}$. □

Lemma C.6. *Let $w = W_p(\mu, \nu)$. Let $y \in E$. Suppose that there exists $x \in \operatorname{supp}(\mu)$ such that $\|x-y\| \leq \alpha$ with $\alpha = (\frac{w}{r^{d-1}})^{\frac{1}{2}}$, and that μ satisfies the Hypotheses 5 and 7 at x with $\frac{1}{2\rho} > r$. Assume that $w \leq (c_5 \wedge 1)(\frac{r}{4})^{d+1}$. Then*

$$W_1(\overline{\mu_y}, \overline{\nu_y}) \leq c_{C.6} r^{\frac{1}{2}} \alpha^{\frac{1}{2}}$$

with $c_{C.6} = \frac{2^{d-2}}{c_5} + \frac{4 \cdot 3^{\frac{1}{2}} 5^{d-\frac{1}{2}} c_7 + 4^{d-\frac{1}{2}}}{3^d c_5} + 2 \cdot 4^d \frac{2c_7(\frac{3}{2})^{d-\frac{1}{2}} + 1}{3^d c_5}$.

Proof. The proof is similar as Lemma C.5. Let us highlight the modifications. Since $\alpha \leq \frac{r}{4}$ and $\frac{w}{\alpha} = r^{d-1}\alpha$, we have the inequalities

$$\alpha^{\frac{1}{2}} \leq \frac{1}{2} r^{\frac{1}{2}}$$
$$\frac{w}{\alpha} \leq \frac{1}{2} r^{d-\frac{1}{2}} \alpha^{\frac{1}{2}}$$

We still write the triangular inequality:

$$W_1(\overline{\mu_y}, \overline{\nu_y}) \leq \underbrace{W_1(\overline{\mu_y}, \overline{p_{1*}\pi_y})}_{(1)} + \underbrace{W_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y})}_{(2)} + \underbrace{W_1(\overline{p_{2*}\pi_y}, \overline{\nu_y})}_{(3)}$$

where π is an optimal transport plan for $W_p(\mu, \nu)$.

Term (2): The argument to obtain $W_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y}) \leq \frac{2^{d-1}}{c_5}\alpha$ is unchanged, and we use $\alpha^{\frac{1}{2}} \leq \frac{1}{2}r^{\frac{1}{2}}$ to get

$$W_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y}) \leq \frac{2^{d-2}}{c_5}\alpha^{\frac{1}{2}}r^{\frac{1}{2}}.$$

Term (1): Using Hypothesis 7, we have

$$\begin{aligned}
&\mu(\overline{\mathcal{B}}(x, r + \|x-y\|)) - \mu(\overline{\mathcal{B}}(x, r - \alpha - \|x-y\|)) \\
&\leq c_7(r + \|x-y\|)^{d-\frac{1}{2}}(2\|x-y\| + \alpha))^{\frac{1}{2}} \\
&\leq c_7\left(\frac{5}{4}r\right)^{d-\frac{1}{2}} 3^{\frac{1}{2}}\alpha^{\frac{1}{2}}.
\end{aligned}$$

And since $\frac{w}{\alpha} \leq \frac{1}{2}r^{d-\frac{1}{2}}\alpha^{\frac{1}{2}}$, we get

$$\begin{aligned}
|\mu_y| - |p_{1*}\pi_y| &\leq \mu(\overline{\mathcal{B}}(x, r + \|x-y\|)) - \mu(\overline{\mathcal{B}}(x, r - \alpha - \|x-y\|)) + \frac{w}{\alpha} \\
&\leq \left(c_7\left(\frac{5}{4}\right)^{d-\frac{1}{2}} 3^{\frac{1}{2}} + \frac{1}{2}\right)r^{d-\frac{1}{2}}\alpha^{\frac{1}{2}}.
\end{aligned}$$

Finally, we use

$$\begin{aligned}
|\mu_y| = \mu(\overline{\mathcal{B}}(y, r)) &\geq \mu(\overline{\mathcal{B}}(x, r - \|x-y\|)) \\
&\geq c_5(r - \|x-y\|)^d \geq c_5\left(\frac{3}{4}\right)^d r^d
\end{aligned}$$

to obtain

$$\frac{|\mu_y| - |p_{1*}\pi_y|}{|\mu_y|} \leq \frac{((c_7(\frac{5}{4})^{d-\frac{1}{2}}3^{\frac{1}{2}} + \frac{1}{2})r^{d-\frac{1}{2}}}{c_5(\frac{3}{4})^d r^d}\alpha^{\frac{1}{2}} = \frac{1}{r^{\frac{1}{2}}} \cdot \frac{2 \cdot 3^{\frac{1}{2}}5^{d-\frac{1}{2}}c_7 + 4^{d-\frac{1}{2}}}{3^d c_5}\alpha^{\frac{1}{2}}$$

and we deduce

$$W_1(\overline{\mu_y}, \overline{p_{1*}\pi_y}) \leq 2\frac{|\mu_y| - |p_{1*}\pi_y|}{|\mu_y|}r \leq \frac{4 \cdot 3^{\frac{1}{2}}5^{d-\frac{1}{2}}c_7 + 4^{d-\frac{1}{2}}}{3^d c_5}r^{\frac{1}{2}}\alpha^{\frac{1}{2}}.$$

Term (3): We use Hypothesis 7 to get

$$\begin{aligned}
&\mu(\overline{\mathcal{B}}(x, r + \|x-y\| + \alpha)) - \mu(\overline{\mathcal{B}}(x, r - \alpha - \|x-y\|)) \\
&\leq c_7(r + \|x-y\| + \alpha)^{d-\frac{1}{2}}(2\|x-y\| + 2\alpha)^{\frac{1}{2}} \\
&\leq 2c_7\left(\frac{3}{2}r\right)^{d-\frac{1}{2}}\alpha^{\frac{1}{2}}.
\end{aligned}$$

And since $\frac{w}{\alpha} \leq \frac{1}{2}r^{d-\frac{1}{2}}\alpha^{\frac{1}{2}}$, we get

$$\begin{aligned}
|\nu_y| - |p_{2*}\pi_y| &\leq \mu(\overline{\mathcal{B}}(x, r + \|x-y\| + \alpha)) - \mu(\overline{\mathcal{B}}(x, r - \alpha - \|x-y\|)) + 2\frac{w}{\alpha} \\
&\leq \left(2c_7\left(\frac{3}{2}\right)^{d-\frac{1}{2}} + 1\right)r^{d-\frac{1}{2}}\alpha^{\frac{1}{2}}.
\end{aligned}$$

Finally, we use

$$\begin{aligned}
|\mu_y| = \mu(\overline{\mathcal{B}}(y, r)) &\geq \mu(\overline{\mathcal{B}}(x, r - \|x-y\|)) \\
&\geq c_5(r - \|x-y\|)^d \geq c_5\left(\frac{3}{4}\right)^d r^d
\end{aligned}$$

to obtain

$$\frac{|\mu_y| - |p_{1*}\pi_y|}{|\mu_y|} \leq \frac{(2c_7(\frac{3}{2})^{d-\frac{1}{2}}+1)r^{d-\frac{1}{2}}}{c_5(\frac{3}{4})^d r^d}\alpha^{\frac{1}{2}} = \frac{1}{r^{\frac{1}{2}}}\cdot 4^d \frac{2c_7(\frac{3}{2})^{d-\frac{1}{2}}+1}{3^d c_5}\alpha^{\frac{1}{2}}$$

and we deduce

$$W_1(\overline{\mu_y}, \overline{p_{1*}\pi_y}) \leq 2\frac{|\mu_y| - |p_{1*}\pi_y|}{|\mu_y|} r \leq 2\cdot 4^d \frac{2c_7(\frac{3}{2})^{d-\frac{1}{2}}+1}{3^d c_5} r^{\frac{1}{2}}\alpha^{\frac{1}{2}}. \qquad \square$$

Remark C.7. Let us comment the inequality of Lemma C.5 with $p = 1$, valid for all r such that $w \leq (a \wedge 1)(\frac{r}{4})^{d+1}$:

$$W_1(\overline{\mu_y}, \overline{\nu_y}) \leq c_{C.5}\left(\frac{w}{r^{d-1}}\right)^{\frac{1}{2}}.$$

If r is assumed to be constant, the behavior of $W_1(\overline{\mu_y}, \overline{\nu_y})$ when w goes to 0 is

$$W_1(\overline{\mu_y}, \overline{\nu_y}) \lesssim w^{\frac{1}{2}}.$$

On the other hand, if r is supposed to follow the worst case, i.e. r is of order $w^{\frac{1}{d+1}}$, then $W_1(\overline{\mu_y}, \overline{\nu_y})$ is of order

$$W_1(\overline{\mu_y}, \overline{\nu_y}) \lesssim \left(\frac{w}{w^{\frac{d-1}{d+1}}}\right)^{\frac{1}{2}} = w^{\frac{1}{d+1}}.$$

Now, let us show that the order $(\frac{w}{r^{d-1}})^{\frac{1}{2}}$ is optimal. More precisely, we show that, for every $d \geq 1$, $r > 0$ and $\epsilon > 0$ fixed, there exists measures μ and ν on $\mathbb{R}^d$ that satisfies the assumptions of Lemma C.5, but such that

$$W_1(\overline{\mu_y}, \overline{\nu_y}) \geq c_d\left(\frac{w}{r^{d-1}}\right)^{\frac{1}{2}} - \epsilon$$

with $c_d = \frac{1}{d+1}\left(\frac{2d}{V_d}\right)^{\frac{1}{2}}$. We consider the following example. Let $\mu = \mathcal{H}^d_{[0,1]^d}$ be the Lebesgue measure on the hypercube $[0,1]^d$. Denote $y = (\frac{1}{2}, ..., \frac{1}{2})$ its center, $B = \mathcal{B}(y, r)$ the open ball, and A the annulus defined as

$$A = \mathcal{B}(y, r+\epsilon) \setminus \mathcal{B}(y, r)$$

where $0 < \epsilon < r < \frac{1}{4}$. In the following, r stays fixed, and ϵ shall go to zero. Consider the probability measure

$$\nu = \mathcal{H}^d_{[0,1]^d \setminus A} + \frac{V_d(r+\epsilon)^d - V_d r^d}{S_{d-1}r^{d-1}}\mathcal{H}^{d-1}_{\partial\mathcal{B}(y,r)}.$$

Let $\overline{\mu_y}$ and $\overline{\nu_y}$ be the localized probability measures associated to μ and ν with parameter r. We shall show that

$$W_1(\mu,\nu) \text{ is of order } r^{d-1}\epsilon^2 \qquad \text{and} \qquad W_1(\overline{\mu_y}, \overline{\nu_y}) \text{ is of order } \epsilon$$

when $\epsilon \to 0$.

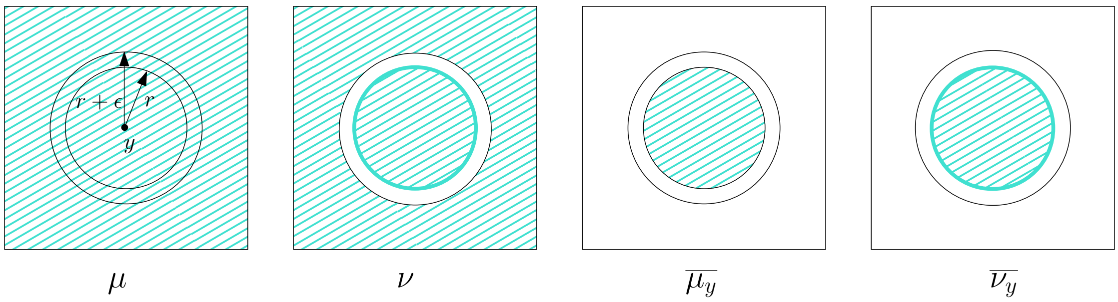

Figure 33: The measures involved in the example. A hatched area represents the d-dimensional Hausdorff measure $\mathcal{H}^d$, and a bold circle represents the $(d-1)$-dimensional Hausdorff measure $\mathcal{H}^{d-1}$.

Step 1: Study of $W_1(\mu,\nu)$. An optimal transport plan between μ and ν is given by transporting the submeasure $\mathcal{H}^d_A$ of μ onto the submeasure $\frac{V_d(r+\epsilon)^d - V_d r^d}{S_{d-1}r^{d-1}}\mathcal{H}^{d-1}_{\partial\mathcal{B}(y,r)}$ of ν via the application

$$\begin{aligned} A &\longrightarrow \partial\mathcal{B}(y,r) \\ x &\longmapsto \frac{r}{\|x\|}x. \end{aligned}$$

Consequently, the Wasserstein distance is

$$W_1(\mu,\nu) = \int_A \left\| x - \frac{r}{\|x\|}x \right\| \frac{V_d(r+\epsilon)^d - V_d r^d}{S_{d-1}r^{d-1}} \mathrm{d}\mathcal{H}^d(x)$$

A change of coordinates shows that

$$\int_A \left\| x - \frac{r}{\|x\|}x \right\| \mathrm{d}\mathcal{H}^d(x) = \int_{\partial\mathcal{B}(0,1)} \int_r^{r+\epsilon} (t-r)t^{d-1}\mathrm{d}\mathcal{H}^1(t)\mathrm{d}\mathcal{H}^{d-1}(v).$$

Let us write $\int_r^{r+\epsilon}(t-r)t^{d-1}\mathrm{d}\mathcal{H}^1(t) = \int_r^{r+\epsilon} t^d \mathrm{d}\mathcal{H}^1(t) - \int_r^{r+\epsilon} rt^{d-1}\mathrm{d}\mathcal{H}^1(t)$. We have

$$\begin{aligned} \int_r^{r+\epsilon} t^d \mathrm{d}\mathcal{H}^1(t) &= \frac{1}{d+1}\left((r+\epsilon)^{d+1} - r^{d+1}\right) \\ &= r^d\epsilon + \frac{d}{2}r^{d-1}\epsilon^2 + o\left(\epsilon^2\right), \end{aligned}$$

where the Little-O notation refers to $\epsilon \to 0$. Moreover,

$$\begin{aligned} \int_r^{r+\epsilon} rt^{d-1}\mathrm{d}\mathcal{H}^1(t) &= r\left(r^{d-1}\epsilon + \frac{d-1}{2}r^{d-2}\epsilon^2 + o\left(\epsilon\right)^2\right) \\ &= r^d\epsilon + \frac{d-1}{2}r^{d-1}\epsilon^2 + o\left(\epsilon^2\right). \end{aligned}$$

We deduce that $\int_r^{r+\epsilon}(t-r)t^{d-1}\mathrm{d}\mathcal{H}^1(t) = \frac{1}{2}r^{d-1}\epsilon^2 + o\left(\epsilon^2\right)$, and

$$\int_A \left\| x - \frac{r}{\|x\|}x \right\| \mathrm{d}\mathcal{H}^d(x) = \frac{S_{d-1}}{2}r^{d-1}\epsilon^2 + o\left(\epsilon^2\right).$$

In other words,

$$W_1(\mu,\nu) = \frac{dV_d}{2}r^{d-1}\epsilon^2 + o\left(\epsilon^2\right).$$

Step 2: Study of $W_1(\overline{\mu_y}, \overline{\nu_y})$. Consider the measures

$$\overline{\mu_x} = \frac{1}{V_d r^d}\mathcal{H}^d_B = \left(\frac{1}{V_d(r+\epsilon)^d} + \frac{V_d(r+\epsilon)^d - V_d r^d}{V_d(r+\epsilon)^d V_d r^d}\right)\mathcal{H}^d_B$$

and

$$\overline{\nu_x} = \frac{1}{V_d(r+\epsilon)^d}\left(\mathcal{H}^d_B + \frac{V_d(r+\epsilon)^d - V_d r^d}{S_{d-1} r^{d-1}}\mathcal{H}^{d-1}_{\partial\mathcal{B}(y,r)}\right).$$

Consider the Wasserstein distance $W_1(\overline{\mu_y}, \overline{\nu_y})$. As before, an optimal transport plan is given by transporting the submeasure $\frac{V_d(r+\epsilon)^d - V_d r^d}{V_d(r+\epsilon)^d V_d r^d}\mathcal{H}^d_B$ of $\overline{\mu_x}$ onto the submeasure $\frac{V_d(r+\epsilon)^d - V_d r^d}{V_d(r+\epsilon)^d S_{d-1} r^{d-1}}\mathcal{H}^{d-1}_{\partial\mathcal{B}(y,r)}$ of $\overline{\nu_x}$. We have:

$$W_1(\overline{\mu_y}, \overline{\nu_y}) = \int_B \left\| x - \frac{r}{\|x\|}x \right\| \frac{V_d(r+\epsilon)^d - V_d r^d}{V_d(r+\epsilon)^d V_d r^d} \mathrm{d}\mathcal{H}^d(x)$$

A change of coordinates yields

$$\int_B \left\| x - \frac{r}{\|x\|}x \right\| = \frac{S_{d-1}}{d(d+1)} r^{d+1}.$$

Besides, we have

$$\frac{V_d(r+\epsilon)^d - V_d r^d}{V_d(r+\epsilon)^d V_d r^d} = \frac{dV_d r^{d-1}\epsilon + O\left(\epsilon^2\right)}{V_d(r+\epsilon)^d V_d r^d} = \frac{d}{V_d}\frac{\epsilon}{r^{d+1}} + O\left(\epsilon^2\right).$$

We deduce that

$$W_1(\overline{\mu_y}, \overline{\nu_y}) = \frac{S_{d-1}}{d(d+1)}\frac{d}{V_d}\epsilon + O\left(\epsilon^2\right) = \frac{d}{d+1}\epsilon + O\left(\epsilon^2\right).$$

Step 3. Using $W_1(\mu,\nu) = \frac{dV_d}{2} r^{d-1}\epsilon^2 + o\left(\epsilon^2\right)$ and $W_1(\overline{\mu_y}, \overline{\nu_y}) = \frac{d}{d+1}\epsilon + O\left(\epsilon^2\right)$, we get

$$\frac{W_1(\overline{\mu_y}, \overline{\nu_y})^2}{W_1(\mu,\nu)} = c\frac{1}{r^{d-1}} + O\left(\epsilon\right)$$

with $c = \frac{\left(\frac{d}{d+1}\right)^2}{\frac{dV_d}{2}} = \frac{2d}{(d+1)^2 V_d}$. In conclusion,

$$\frac{W_1(\overline{\mu_y}, \overline{\nu_y})}{W_1(\mu,\nu)^{\frac{1}{2}}} = c^{\frac{1}{2}}\left(\frac{1}{r^{d-1}}\right)^{\frac{1}{2}} + O\left(\epsilon\right),$$

and since $W_1(\mu,\nu)^{\frac{1}{2}} = O\left(\epsilon\right)$, we deduce

$$W_1(\overline{\mu_y}, \overline{\nu_y}) = c^{\frac{1}{2}}\left(\frac{W_1(\mu,\nu)}{r^{d-1}}\right)^{\frac{1}{2}} + O\left(\epsilon^2\right).$$

References

[Aam18] Eddie Aamari. *Vitesses de convergence en inférence géométrique.* PhD thesis, Université Paris-Saclay, 2018.

[AB06] Stephanie B Alexander and Richard L Bishop. Gauss equation and injectivity radii for subspaces in spaces of curvature bounded above. *Geometriae Dedicata*, 117(1):65–84, 2006.

[ACG$^+$18] Hirokazu Anai, Frédéric Chazal, Marc Glisse, Yuichi Ike, Hiroya Inakoshi, Raphaël Tinarrage, and Yuhei Umeda. DTM-based filtrations. In *Symposium on Computational Geometry*, 2018.

[ACLZ17] Ery Arias-Castro, Gilad Lerman, and Teng Zhang. Spectral clustering based on local PCA. *The Journal of Machine Learning Research*, 18(1):253–309, 2017.

[AKC$^+$19] Eddie Aamari, Jisu Kim, Frédéric Chazal, Bertrand Michel, Alessandro Rinaldo, and Larry Wasserman. Estimating the Reach of a Manifold. *Electronic journal of statistics* , April 2019.

[BLW19] Jean-Daniel Boissonnat, André Lieutier, and Mathijs Wintraecken. The reach, metric distortion, geodesic convexity and the variation of tangent spaces. *Journal of Applied and Computational Topology*, 3(1-2):29–58, 2019.

[CCSL09] Frédéric Chazal, David Cohen-Steiner, and André Lieutier. A sampling theory for compact sets in euclidean space. *Discrete & Computational Geometry*, 41(3):461–479, 2009.

[CCSM11] F. Chazal, D. Cohen-Steiner, and Q. Mérigot. Geometric inference for probability measures. *Journal on Found. of Comp. Mathematics*, 11(6):733–751, 2011.

[CdSGO16] Frédéric Chazal, Vin de Silva, Marc Glisse, and Steve Oudot. *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics, 2016.

[dC92] M.P. do Carmo. *Riemannian Geometry*. Mathematics (Boston, Mass.). Birkhäuser, 1992.

[DVW15] Ramsay Dyer, Gert Vegter, and Mathijs Wintraecken. Riemannian simplices and triangulations. *Geometriae Dedicata*, 179(1):91–138, 2015.

[Fed59] Herbert Federer. Curvature measures. *Transactions of the American Mathematical Society*, 93(3):418–491, 1959.

[Fed14] Herbert Federer. *Geometric measure theory*. Springer, 2014.

[MMM18] Diego H Díaz Martínez, Facundo Mémoli, and Washington Mio. The shape of data and probability measures. *Applied and Computational Harmonic Analysis*, 2018.

[MSW19] Facundo Memoli, Zane Smith, and Zhengchao Wan. The Wasserstein transform. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 4496–4504, Long Beach, California, USA, 09–15 Jun 2019. PMLR.

[NSW08] Partha Niyogi, Stephen Smale, and Shmuel Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete & Computational Geometry*, 39(1-3):419–441, 2008.